



TUGAS AKHIR - SS141501

**PEMODELAN HASIL KEMOTERAPI BERBASIS
5-FLUOROURACIL MENGGUNAKAN METODE
REGRESI MULTIVARIAT DAN KLASIFIKASI
MENGGUNAKAN REGRESI LOGISTIK BINER**

**TAUFIK AFIIF MALDINI
NRP 062114 4000 0016**

**Dosen Pembimbing
Dr.rer.pol. Heri Kuswanto, S.Si., M.Si.**

**PROGRAM STUDI SARJANA
DEPARTEMEN STATISTIKA
FAKULTAS MATEMATIKA, KOMPUTASI, DAN SAINS DATA
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
SURABAYA 2018**



TUGAS AKHIR - SS141501

**PEMODELAN HASIL KEMOTERAPI BERBASIS
5-FLUOROURACIL MENGGUNAKAN METODE
REGRESI MULTIVARIAT DAN KLASIFIKASI
MENGGUNAKAN REGRESI LOGISTIK BINER**

**TAUFIK AFIIF MALDINI
NRP 062114 4000 0016**

**Dosen Pembimbing
Dr.rer.pol Heri Kuswanto, S.Si., M.Si.**

**PROGRAM STUDI SARJANA
DEPARTEMEN STATISTIKA
FAKULTAS MATEMATIKA, KOMPUTASI, DAN SAINS DATA
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
SURABAYA 2018**



FINAL PROJECT - SS 141501

**MODELLING OF CHEMOTHERAPY TREATMENT
RESULT BASED ON *5-FLUOROURACIL*
COMPOUND USING MULTIVARIATE REGRESSION
METHOD AND CLASSIFICATION USING BINARY
LOGISTIC REGRESSION**

**TAUFIK AFIIF MALDINI
SN 062114 4000 0016**

**Supervisor
Dr.rer.pol Heri Kuswanto, S.Si., M.Si.**

**UNDERGRADUATE PROGRAMME
DEPARTMENT OF STATISTICS
FACULTY OF MATHEMATICS, COMPUTING, AND DATA SCIENCE
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
SURABAYA 2018**

LEMBAR PENGESAHAN

PEMODELAN HASIL KEMOTERAPI BERBASIS 5-FLUOROURACIL MENGGUNAKAN METODE REGRESI MULTIVARIAT DAN KLASIFIKASI MENGGUNAKAN REGRESI LOGISTIK BINER

TUGAS AKHIR

Diajukan Untuk Memenuhi Salah Satu Syarat
Memperoleh Gelar Sarjana Sains
pada

Program Studi Sarjana Departemen Statistika
Fakultas Matematika, Komputasi, dan Sains Data
Institut Teknologi Sepuluh Nopember

Oleh :

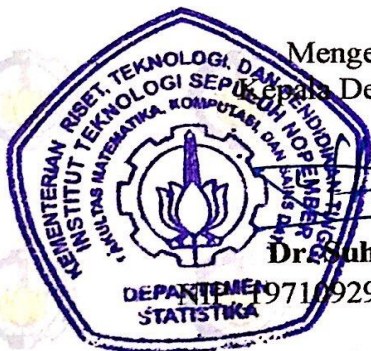
Taufik Afiif Maldini

NRP. 062114 4000 0016

Disetujui oleh Pembimbing:

Dr.rer.pol Heri Kuswanto, S.Si, M.Si

NIP. 19820326 200312 1 004



Mengetahui,
Kepala Departemen

Dr. Suhartono

NIP. 19710929199512 1 001

SURABAYA, JULI 2018

**PEMODELAN HASIL KEMOTERAPI BERBASIS
5-FLUOROURACIL MENGGUNAKAN METODE REGRESI
MULTIVARIAT DAN KLASIFIKASI MENGGUNAKAN
REGRESI LOGISTIK BINER**

Nama Mahasiswa : Taufik Afif Maldini
NRP : 06114 4000 0016
Departemen : Statistika-FMKSD-ITS
Dosen Pembimbing : Dr.rer.pol Heri Kuswanto, S.Si., M.Si.

Abstrak

Kanker adalah suatu istilah untuk penyakit di mana sel-sel membelah secara abnormal tanpa kontrol dan dapat menyerang jaringan di sekitarnya. Dalam beberapa tahun terakhir, kanker telah menjadi penyebab utama kematian di dunia. Pengobatan kanker yang paling umum digunakan adalah kemoterapi dengan rangkaian terapi berbasis 5-Fluorouracil (5-FU). Pada hasil analisis regresi multivariat didapatkan bahwa tidak terdapat variabel prediktor yang berpengaruh dalam perubahan jumlah White Blood Cells (WBC), perubahan jumlah neutrophil dan rata-rata jam tidur tidak nyenyak pasien setelah melakukan kemoterapi. Menggunakan regresi linear berganda didapatkan hasil bahwa dosis senyawa 5-FU 1, Irinotecan(IRI) dan Indeks Massa Tubuh pasien berpengaruh signifikan pada perubahan WBC dengan taraf signifikansi 20%. Pada perubahan neutrophil pasien dipengaruhi oleh dosis senyawa Irinotecan (IRI) dan rata-rata jam tidur tidak nyenyak pasien dipengaruhi oleh usia pasien dan senyawa P (Panitumumab) dengan taraf signifikansi 10%. Pada analisis regresi logistik didapatkan hasil bahwa usia pasien, rata-rata jumlah langkah kaki dan jumlah dosis 5 FU-1 yang diberikan pada saat kemoterapi berpengaruh signifikan pada perubahan WBC setelah kemoterapi dengan taraf signifikansi 10% dengan AUC sebesar 75,44%.

Kata Kunci: Fluorouracil, Kanker, Kemoterapi, Regresi Logistik, Regresi Multivariat

(Halaman ini sengaja dikosongkan)

**MODELLING OF CHEMOTHERAPY TREATMENT
RESULT BASED 5-FLUOROURACIL COMPOUND USING
MULTIVARIATE REGRESSION METHOD AND
CLASSIFICATION USING BINARY LOGISTIC
REGRESSION**

Name : Taufik Afiif Maldini
Student Number : 062114 4000 0016
Department : Statistics-FMKSD-ITS
Supervisor : Dr.rer.pol Heri Kuswanto, S.Si., M.Si.

Abstract

Cancer is a term for a disease in which cells divide abnormally without control and can invade surrounding tissue. In recent years, cancer has been the leading cause of death in the world. The most commonly used cancer treatment is chemotherapy with a 5-Fluorouracil (5-FU) based therapeutic regimen. In the multivariate regression analysis it was found that there were no predictor variabels that influenced the number of White Blood Cells (WBC), the change of neutrophil count and the average hours of unperturbed sleep after chemotherapy. Using multiple linear regression, it was found that the dose of 5-FU 1 compound, Irinotecan (IRI) and patient body mass index had significant effect on WBC changes on the level of 20% significance. In the neutrophil changes the patient is affected by the dose of Irinotecan (IRI) compound and the average unbearable sleep time of the patient is affected by the patient's age and the P (Panitumumab) compound with a significance level of 10%. In the logistic regression analysis the results showed that the age of the patients, the average number of foot steps and the number of doses of 5 FU-1 given at chemotherapy had a significant effect on WBC changes after chemotherapy with a significance level of 10% with AUC of 75.44%.

Keywords: *Fluorouracil, Cancer, Chemotherapy, Logistic Regression, Multivariate Regression*

(Halaman ini sengaja dikosongkan)

KATA PENGANTAR

Puji syukur penulis panjatkan atas rahmat dan hidayah yang diberikan Allah SWT sehingga penulis dapat menyelesaikan laporan Tugas Akhir yang berjudul **“Pemodelan Hasil Kemoterapi Berbasis 5-Fluorouracil Menggunakan Metode Regresi Multivariat dan Klasifikasi Menggunakan Regresi Logistik Biner”** dengan lancar.

Penulis menyadari bahwa Tugas Akhir ini dapat terselesaikan tidak terlepas dari bantuan dan dukungan berbagai pihak. Oleh karena itu, penulis menyampaikan terima kasih kepada :

1. Kedua orang tua dan keluarga atas segala do'a, nasihat, kasih sayang, dan dukungan yang diberikan kepada penulis demi kesuksesan dan kebahagiaan penulis.
2. Bapak Dr. Suhartono selaku Kepala Departemen Statistika Fakultas Matematika, Komputasi dan Sains Data ITS.
3. Dr. Sutikno, M.Si. selaku Ketua Program Studi Sarjana yang telah memberikan fasilitas, sarana, dan prasarana.
4. Dr.rer.pol Heri Kuswanto, S.Si., M.Si. selaku dosen pembimbing Tugas Akhir yang telah meluangkan waktu dan dengan sangat sabar memberikan bimbingan, saran, dukungan serta motivasi selama penyusunan Tugas Akhir.
5. Dr. Ir. Setiawan, M.S. dan M. Sjahid Akbar, S.Si., M.Si. selaku dosen penguji yang telah banyak memberi masukan kepada penulis.
6. Prof. Drs. NUR Iriawan, M.Ikomp, Ph.D selaku dosen wali yang telah banyak memberikan saran dan arahan dalam proses belajar di Departemen Statistika.
7. Seluruh dosen dan *staff* pengajar Program Studi Sarjana Departemen Statistika FMKSD ITS yang telah membantu penulis selama di bangku kuliah
8. Teman-teman Statistika ITS angkatan 2014, Respect, yang selalu memberikan dukungan kepada penulis selama ini.
9. Tim bimbingan Tugas Akhir Pak Heri yang selalu memeberikan informasi dan bantuan dalam menyelesaikan Tugas Akhir ini.

10. Teman-teman Garuda, Gejo, Wacana dan Trondols yang selalu memberikan dukungan, motivasi serta semangatnya dalam penyelesaian Tugas Akhir ini
11. Teman-teman IASMA MUDA Surabaya dan IMAMI Surabaya selaku keluarga pertama di Surabaya yang sama-sama berjuang di negeri perantauan.
12. Teman-teman seperjuangan TA demi Toga 118 atas perjuangan dan kebersamaannya menyelesaikan Tugas Akhir ini
13. Semua pihak yang turut membantu dalam pelaksanaan Tugas Akhir yang tidak bisa penulis sebutkan satu persatu.

Besar harapan penulis untuk mendapatkan kritik dan saran yang membangun sehingga Tugas Akhir ini dapat memberikan manfaat bagi semua pihak yang terkait.

Surabaya, Juli 2018

Penulis

DAFTAR ISI

HALAMAN JUDUL	i
LEMBAR PENGESAHAN	v
ABSTRAK	vii
ABSTRACT	ix
KATA PENGANTAR	xi
DAFTAR ISI	xiii
DAFTAR GAMBAR	xvii
DAFTAR TABEL	xix
DAFTAR LAMPIRAN	xxi
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	6
1.3 Tujuan	6
1.4 Manfaat	7
BAB II TINJAUAN PUSTAKA	9
2.1 Analisis Regresi Multivariat	9
2.1.1 Uji Independensi	10
2.1.2 Uji Distribusi Normal Multivariat	10
2.1.3 Estimasi Parameter	12
2.1.4 Pengujian Signifikansi Parameter Model	13
2.1.5 Uji Asumsi Residual IIDN	14
2.1.6 Ukuran Kebaikan Model Regresi Multivariat	17
2.2 Regresi Linear Berganda	18
2.2.1 Pengujian Hipotesis Model Regresi Linear Berganda	19
2.2.2 Pengujian Asumsi Residual	20
2.2.3 Ukuran Kebaikan Model Regresi Linear	23
2.3 Regresi Logistik Biner	23
2.3.1 Estimasi Parameter	25
2.3.2 Pengujian Signifikansi Parameter	26
2.3.3 Interpretasi Model	27
2.3.4 Evaluasi dan Validasi Klasifikasi Model	28
2.4 Tinjauan Non Statistik	29
2.4.1 Kanker	29

2.4.3	Kemoterapi	30
2.4.4	5-Fluorouracil (5-FU)	31
2.4.5	White Blood Cells (WBC).....	32
2.4.6	Neutrophil.....	32
2.4.7	Indeks Massa Tubuh (IMT)	34
BAB III	METODOLOGI PENELITIAN	35
3.1	Sumber Data	35
3.2	Variabel Penelitian	35
3.3	Struktur Data	37
3.4	Langkah Penelitian	37
BAB IV	ANALISIS DAN PEMBAHASAN	43
4.1	Karakteristik data pasien kemoterapi kanker	43
4.2	Pemodelan Menggunakan Regresi Multivariat	47
4.2.1	Pengujian Dependensi Antar Variabel Respon.....	53
4.2.2	Pengujian Normal Multivariat	54
4.2.3	Estimasi Parameter	55
4.2.4	Pengujian Signifikansi Serentak Secara Multivariat	55
4.2.5	Pengujian Signifikansi Parsial Secara Multivariat.....	55
4.3	Pemodelan Menggunakan Regresi Linear Berganda Menggunakan Metode <i>Stepwise</i>	58
4.3.1	Model Regresi Linear Berganda pada Perubahan WBC	58
4.3.2	Pengujian Asumsi Model Regresi Linear Perubahan WBC	60
4.3.3	Model Regresi Linear Berganda pada Perubahan <i>Neutrophil</i>	61
4.3.4	Pengujian Asumsi Model Regresi Linear Perubahan <i>Neutrophil</i>	62
4.3.5	Model Regresi Linear Berganda pada Jam Tidur Tidak Nyenyak Pasien	63
4.3.6	Pengujian Asumsi Model Regresi Linear Jam tidur tidak nyenyak Pasien.....	65

4.4	Pemodelan Menggunakan Analisis Regresi Logistik..	66
4.4.1	Karakteristik Variabel yang Digunakan pada Analisis Regresi Logistik Biner	66
4.4.2	Estimasi Parameter Menggunakan Metode <i>Wald</i>	67
4.4.3	Ketepatan Klasifikasi.....	68
4.4.4	Hasil Analisis Regresi Logistik Biner.....	69
BAB V	KESIMPULAN DAN SARAN.....	71
5.1	Kesimpulan.....	71
5.2	Saran.....	72
DAFTAR PUSTAKA	74
LAMPIRAN	80
BIODATA PENULIS	113

(Halaman ini sengaja dikosongkan)

DAFTAR GAMBAR

Gambar 3.1	Diagram Alir regresi Multivariat.....	40
Gambar 3.2	Diagram Alir Regresi Logistik Biner	41
Gambar 4.1	Proporsi Pasien Kemoterapi Kanker Berdasarkan Jenis Kelamin	45
Gambar 4.2	Proporsi Pasien Kemoterapi Kanker Berdasarkan Indeks Massa Tubuh.....	45
Gambar 4.3	Proporsi Pasien Menggunakan Senyawa Lain Saat Kemoterapi	46
Gambar 4.4	Hubungan Variabel Prediktor Terhadap Variabel Y_1	48
Gambar 4.5	Hubungan Variabel Prediktor Terhadap Variabel Y_2	49
Gambar 4.6	Hubungan Variabel Prediktor Terhadap Variabel Y_3	50
Gambar 4.7	Variabel Respon Berdasarkan Jenis Kelamin	52
Gambar 4.8	Variabel Respon Berdasarkan Indeks Massa Tubuh	53
Gambar 4.9	QQ Plot Variabel Respon	54
Gambar 4.10	Proporsi Klasifikasi Perubahan WBC Pada Pasien	66

(Halaman ini sengaja dikosongkan)

DAFTAR TABEL

Tabel 2.1	Tabel Klasifikasi.....	28
Tabel 2.2	Klasifikasi IMT	33
Tabel 3.1	Variabel Respon Penelitian.....	35
Tabel 3.2	Variabel Prediktor Penelitian.....	36
Tabel 3.3	Struktur Data Regresi Multivariat.....	37
Tabel 3.4	Struktur Data Regresi Logistik Biner	38
Tabel 4.1	Statistika Deskriptif Variabel Respon.....	43
Tabel 4.2	Statistika Deskriptif Variabel Prediktor	44
Tabel 4.3	Korelasi Antar Variabel.....	51
Tabel 4.4	Pengujian Signifikansi Parsial Secara Multivariat	57
Tabel 4.5	<i>Analysis of Variance</i> (ANOVA) Variabel Y_1	58
Tabel 4.6	Estimasi dan Pengujian Parameter Model Regresi Linear Y_1	59
Tabel 4.7	Persamaan regresi linear Y_1	59
Tabel 4.8	<i>Analysis of Variance</i> (ANOVA) Variabel Y_2	61
Tabel 4.9	Estimasi dan Pengujian Parameter Model Regresi Linear Y_2	62
Tabel 4.10	<i>Analysis of Variance</i> (ANOVA) Variabel Y_3	63
Tabel 4.11	Estimasi dan Pengujian Parameter Model Regresi Linear Y_3	64
Tabel 4.12	Seleksi Model Terbaik	67
Tabel 4.13	Hasil Prediksi Klasifikasi WBC	68

(Halaman ini sengaja dikosongkan)

DAFTAR LAMPIRAN

Lampiran 1	Data Pasien Kemoterapi Kanaker Berbasis 5-FU .	81
Lampiran 2	<i>Output</i> Minitab Korelasi Antar Variabel	84
Lampiran 3	Macro Minitab untuk pengujian Normal Multivariat	87
Lampiran 4	<i>Output</i> SPSS Uji Independensi Variabel Respon.....	88
Lampiran 5	<i>Output</i> SPSS Model Regresi Multivariat.....	89
Lampiran 6	<i>Output</i> SPSS Pengujian Signifikansi Multivariat Model	92
Lampiran 7	<i>Output</i> Minitab Pengujian Signifikansi Parameter Secara Serentak pada Pemodelan Y_1	96
Lampiran 8	<i>Output</i> Minitab Pengujian Signifikansi Parameter Secara Parsial pada Pemodelan Y_1	96
Lampiran 9	Koefisien determinasi (R^2) Model regresi linear Y_1	96
Lampiran 10	<i>Output</i> Minitab asumsi Residual IIDN Model pada Variabel Y_1	97
Lampiran 11	<i>Output</i> Minitab Pengujian Signifikansi Parameter Secara Serentak pada Pemodelan Y_2	99
Lampiran 12	<i>Output</i> Minitab Pengujian Signifikansi Parameter Secara Parsial pada Pemodelan Y_2	99
Lampiran 13	Koefisien determinasi (R^2) Model regresi linear Y_2	99
Lampiran 14	<i>Output</i> Minitab asumsi Residual IIDN Model pada Variabel Y_2	100
Lampiran 15	<i>Output</i> Minitab Pengujian Signifikansi Parameter Secara Serentak pada Pemodelan Y_3	102
Lampiran 16	<i>Output</i> Minitab Pengujian Signifikansi Parameter Secara Parsial pada Pemodelan Y_3	102

Lampiran 17	Koefisien determinasi (R^2) Model regresi linear Y_3	102
Lampiran 18	<i>Output</i> Minitab asumsi Residual IIDN Model pada Variabel Y_3	103
Lampiran 19	<i>Output</i> SPSS Model Regresi Logistik Biner Menggunakan Metode <i>Stepwiset</i>	105
Lampiran 20	<i>Output</i> SPSS Ketepatan Klasifikasi	111
Lampiran 21	Surat Pernyataan.....	112

BAB I

PENDAHULUAN

1.1 Latar Belakang

Menurut Yayasan Kanker Indonesia (2018), Kanker adalah penyakit akibat pertumbuhan tidak normal dari sel-sel jaringan tubuh yang berubah menjadi sel kanker. Dalam perkembangannya, sel-sel kanker ini dapat menyebar ke bagian tubuh lainnya sehingga dapat menyebabkan kematian. Kanker adalah pertumbuhan sel-sel baru secara abnormal yang tumbuh melampaui batas normal, dan yang kemudian dapat menyerang bagian sebelah tubuh dan menyebar ke organ lain. Menurut National Cancer Institute, kanker adalah suatu istilah untuk penyakit di mana sel-sel membelah secara abnormal tanpa kontrol dan dapat menyerang jaringan di sekitarnya. Proses ini disebut metastasis, yang merupakan penyebab utama kematian akibat kanker (World Health Organisation, 2018). Kanker terjadi karena kerusakan struktur genetik yang menyebabkan pertumbuhan sel menjadi tidak terkontrol. Pola insiden kanker bervariasi sesuai jenis kelamin, ras, dan letak geografik. Beberapa kanker dapat dipengaruhi oleh faktor genetik keluarga, namun yang paling sering terjadi karena faktor lingkungan dan gaya hidup (Wahyuni, 2015).

Dalam beberapa tahun terakhir, kanker adalah penyebab utama tingkat kematian dan angka kematian di seluruh dunia. Menurut WHO, Kanker adalah penyebab utama kematian kedua di dunia dan menyumbang 8,8 juta kematian pada tahun 2015. Berdasarkan Data *Global Burden of Cancer, International Agency for Research on Cancer (IARC)*, diketahui bahwa pada tahun 2012 kasus kanker mengakibatkan 8.2 juta kematian (58.3%) dari total 14.1 juta kasus baru kanker di seluruh dunia. Di Jepang sendiri, Kasus kanker juga merupakan penyebab utama kematian di negara tersebut. Selama beberapa tahun setelah Perang Dunia II, antara 50.000 dan 60.000 orang di Jepang meninggal karena kanker setiap tahunnya. Sejak saat itu, jumlah kematian akibat kanker terus meningkat. Menurut statistik yang dikumpulkan oleh Kementerian Kesehatan, Tenaga Kerja dan Kesejahteraan di Jepang sekitar

353.000 orang meninggal karena kanker pada tahun 2010, atau sekitar satu dari tiga kasus kematian di Jepang (Japan Cancer Society, 2011). *Union for International Cancer Control* (UICC) memprediksi pada tahun 2030, jumlah penderita kanker akan terus mengalami peningkatan sebesar 300% di seluruh dunia (Kartika, 2013).

Pengobatan yang dilakukan pada penderita kanker dilakukan dengan berbagai metode pengobatan seperti operasi, kemoterapi, imunoterapi sel, dll. Pilihan pengobatan yang paling umum pada kanker adalah bedah. Namun, Tergantung pada stadium kanker, kemoterapi mungkin diperlukan setelah operasi untuk meningkatkan kemungkinan seseorang untuk sembuh dari kanker. Kemoterapi atau biasa disebut kemo dikenal sebagai pengobatan untuk membunuh sel kanker (National Cancer Institut, 2015). Kemoterapi dilakukan melalui suntikan obat anti-kanker ke dalam pembuluh darah di tangan. Terapi ini mengandalkan kemampuan dari obat-obat khusus untuk menghancurkan sel-sel kanker yang menyerang tubuh. Obat tersebut bekerja dengan memperlambat maupun menghentikan pertumbuhan sel kanker. komponen obat yang banyak digunakan dalam pengobatan kanker adalah *5-Fluorouracil* (Longley & Johnston, 2007).

Saat ini standar emas dalam mengobati pasien penderita kanker adalah kemoterapi dengan rangkaian terapi berbasis *5-Fluorouracil* (*5-FU*). Penelitian telah dilakukan oleh Moertel et al., (1995) mengenai keefektifan kemoterapi menggunakan *5-FU* dan *levamisole* dalam meningkatkan tingkat penyembuhan pada penderita kanker usus stadium 3. Pada penelitian tersebut diukur tingkat kekambuhan kanker dan kematian serta efek samping awal dan akhir pengobatan. Diperoleh hasil bahwa kemoterapi *Fluorouracil plus levamisole* menurunkan tingkat kekambuhan sebesar 40% dan tingkat kematian sebesar 33%. *Levamisole* mengurangi tingkat kekambuhan hanya 2% dan tingkat kematian hanya 6%. dan tidak ada bukti efek samping yang terlihat. Disimpulkan bahwa *5-FU* dan *levamisole* adalah terapi penyembuhan yang sesuai setelah operasi, dan banyak

meningkatkan angka penyembuhan bagi pasien penderita kanker kolon risiko tinggi (stadium 3). Selain itu, Kohne, Cunningham et al., (2002) melakukan penelitian mengenai Faktor penentu klinis untuk bertahan hidup pada pasien dengan perawatan berbasis *5-Fluorouracil* untuk kanker kolorektal metastatik. Penelitian lain juga dilakukan Giacchetti, Itzhaki et al., (1999) mengenai *Long-term survival* pada pasien metastasis kanker kolorektal yang tidak dapat diobati melalui operasi, lalu mencoba kemoterapi dengan 5-fluorourasil, leucovorin, oxaliplatin dan pembedahan.

Penelitian menggunakan Regresi Multivariat dalam kasus kanker pernah dilakukan oleh Chapuis et al., (1985). Dimana pada penelitian tersebut dilakukan analisis multivariat pada variabel klinis dan patologis dalam prognosis setelah reseksi kanker usus besar. Terdapat beberapa variabel yang berpengaruh yaitu tingkat histologis, tingkat penyebaran langsung, adanya invasi vena, usia dan jenis kelamin pasien dan adanya penyumbatan. Penelitian lain oleh Peng et al., (2014). Pada penelitian tersebut, mencoba sebuah metode baru remMap (Regresi multivarian teregregasi untuk mengidentifikasi Prediktor Master) untuk menyesuaikan model regresi respons multivariat di bawah pengaturan ukuran sampel dengan dimensi rendah. Penelitian oleh Setiarni et al., (2011) mengenai hubungan antara tingkat pengetahuan, status ekonomi dan kebiasaan merokok dengan kejadian tuberkulosis paru pada orang dewasa di wilayah kerja puskesmas tuan-tuan kabupaten ketapang kalimantan barat. Analisis data pada penelitian tersebut dilakukan dengan distribusi frekuensi, analisis bivariat menggunakan analisis *Chi-square* dan multivariat dengan regresi logistik. Hasil penelitian menunjukkan bahwa adanya hubungan variabel tingkat pengetahuan dan kebiasaan merokok dengan kasus tuberkulosis paru pada orang dewasa di wilayah Puskesmas Tuan-Tuan Ketapang. Dari hasil analisis multivariat variabel yang paling dominan terkait kasus tuberkulosis paru pada orang dewasa di wilayah Puskesmas Tuan-Tuan Ketapang adalah kebiasaan merokok dengan nilai signifikan 0,012

Penelitian mengenai kasus pemberian *5-FU* pada kemoterapi telah dilakukan oleh Tamoki pada tahun 2017. Dimana pada penelitian tersebut dibahas mengenai ekstraksi efek samping yang diberikan setelah melakukan kemoterapi berbasis *5-FU*. Penelitian ini merupakan penelitian yang relatif baru di dunia kemoterapi. Dalam penelitian ini mencoba untuk menganalisis lebih jauh tentang efek kemoterapi yang dikombinasikan dengan *activity meter* yang dipasangkan pada tubuh pasien setelah melakukan kemoterapi. Selama 14 hari setelah menjalani kemoterapi, *activity meter* digunakan untuk mengukur jumlah langkah kaki dan jam tidur tidak nyenyak pada pasien. Berbeda dengan penelitian yang dilakukan oleh Tamoki yang menggunakan metode machine learning, penelitian kali ini memiliki tujuan yang berbeda. Penelitian ini dilakukan untuk diketahui faktor-faktor apa yang dapat mempengaruhi keberhasilan kemoterapi dengan rangkaian terapi berbasis *5-FU*.

Pasien kanker perlu mempertimbangkan terapi dengan cermat sebelum mereka menjalani tindakan tersebut. Ini karena kemoterapi juga dapat membahayakan sel-sel sehat yang membagi diri dengan cepat, tidak hanya sel ganas. terdapat banyak efek samping dari kemoterapi, Efek samping dari *5-FU* yang ditemukan pada pasien antara lain neutropenia, stomatitis, diare, dan *hand-food syndrome*. Masing-masing efek ini terkait dengan metode pemberian yang diterapkan pada pasien (Meyerhardt & Mayer, 2005). Setelah melakukan pengobatan kemoterapi pada pasien, penilaian terhadap perubahan kanker merupakan fitur penting sebagai evaluasi klinis terapi untuk dapat melakukan tindakan pengobatan selanjutnya. Tindakan yang tepat pada pasien dilakukan dengan mempertimbangkan faktor-faktor yang mempengaruhi tingkat keberhasilan pemberian kemoterapi pada pasien agar mengurangi efek sampingnya.

Berdasarkan penjelasan yang telah diulas sebelumnya, maka pada penelitian ini untuk mengetahui faktor-faktor yang dianggap berpengaruh terhadap keberhasilan kemoterapi, dapat dilihat berdasarkan pengaruh faktor tersebut dalam perubahan jumlah

neutrophil, perubahan *White Blood Cells* (WBC), dan rata-rata jam tidur tidak nyenyak pasien setelah kemoterapi menggunakan metode regresi multivariat. Metode regresi multivariat lebih tepat digunakan daripada regresi biasa atau regresi yang lainnya karena model regresi multivariat adalah model regresi dengan lebih dari satu variabel respon yang saling berkorelasi dan satu atau lebih variabel prediktor (Johnson dan Wichern, 2007).

Selain itu, dalam penelitian ini juga akan diklasifikasikan salah satu efek samping dari kemoterapi yaitu perubahan *White Blood Cells* (WBC) pada pasien setelah melakukan kemoterapi, yaitu ketika jumlah *White Blood Cells* pasien mengalami peningkatan setelah melakukan kemoterapi, maka kemoterapi dapat dikategorikan tidak terjadi efek samping dan jika *White Blood Cells* pasien mengalami penurunan jumlah setelah melakukan kemoterapi maka dapat dikategorikan terjadi efek samping pada pasien setelah melakukan kemoterapi. Jumlah sel darah putih (WBC) yang rendah adalah salah satu efek samping yang cukup serius dari kemoterapi. Sel darah putih diperlukan bagi tubuh sebagai pertahanan untuk melawan infeksi. Resiko infeksi akan lebih tinggi ketika jumlah sel darah putih (WBC) rendah.

Faktor-faktor yang mempengaruhi perubahan WBC tersebut akan untuk mengetahuinya dapat dianalisis menggunakan metode regresi logistik. Regresi Logistik merupakan metode statistik yang bertujuan untuk mengetahui hubungan variabel respon (kategorik) dengan satu atau lebih variabel prediktor yang bersifat kategorik/kontinu (Agresti, 2002). Ada beberapa kelebihan klasifikasi Regresi Logistik dibandingkan metode klasifikasi lainnya yang telah dipelajari. Regresi Logistik memberikan probabilitas dan mencakup masalah klasifikasi multi kelas sehingga disebut sebagai *classifier* yang tangguh (Hastie, Tibshirani, & Friedman, 2001). Keuntungan menggunakan Regresi Logistik adalah metode ini telah dipelajari secara ekstensif, telah diperbaiki dengan menggunakan metode Newton yang dipersingkat, dan tidak memiliki asumsi untuk distribusi pada variabel independen (Maalouf & Trafalis, 2010).

1.2 Rumusan Masalah

Tingginya kasus kematian akibat kanker dan prediksi jumlah penderita kanker yang akan terus mengalami peningkatan, membuat dunia kesehatan harus selalu mengevaluasi kinerja medis yang berkaitan erat dengan risiko kanker yang seharusnya dapat dicegah. Metode penyembuhan seperti operasi (pembedahan), radioterapi (menggunakan sinar *gamma-ray*), imunoterapi, dan kemoterapi adalah metode yang sering digunakan. Kemoterapi adalah metode yang paling sering digunakan oleh pasien karena kemoterapi terbukti dapat mematikan sel-sel kanker dan mengecilkan ukuran tumor. Namun, terdapat dampak dan efek samping dari penggunaan kemoterapi. Oleh karena itu, perlu diketahui faktor-faktor yang mempengaruhi keberhasilan kemoterapi kanker agar mengurangi efek samping tersebut. Faktor-faktor yang diduga mempengaruhi hasil kemoterapi kanker dapat dimodelkan menggunakan Analisis Regresi Multivariat karena pada penelitian ini variabel respon yang digunakan berjumlah lebih dari satu. Selain itu juga ingin diketahui faktor-faktor yang mempengaruhi hasil kemoterapi berbasis *5-FU* dengan mencari model klasifikasinya. Hasil klasifikasi keberhasilan kemoterapi kanker akan diamati menggunakan metode regresi logistik.

1.3 Tujuan

Berdasarkan rumusan masalah yang telah diuraikan di atas, tujuan yang ingin dicapai dalam penelitian ini adalah sebagai berikut.

1. Mendeskripsikan karakteristik data pasien yang melakukan kemoterapi menggunakan komponen *5-FU*.
2. Menentukan faktor-faktor yang dianggap berpengaruh pada perubahan jumlah *White Blood Cells* (WBC), perubahan jumlah *Neutrophil* dan rata-rata jam tidur tidak nyenyak pasien selama 14 hari setelah melakukan kemoterapi.
3. Mengklasifikasikan keberhasilan kemoterapi kanker yang dilihat berdasarkan perubahan *White Blood Cells* (WBC) pasien setelah melakukan kemoterapi serta mengetahui faktor-faktor yang mempengaruhinya.

1.4 Manfaat

Pada penelitian ini dengan penerapan ilmu statistika, dapat diketahui karakteristik pasien setelah melakukan kemoterapi, faktor-faktor yang mempengaruhi keberhasilan suatu kemoterapi, serta hasil klasifikasi keberhasilan kemoterapi pada pasien. Diharapkan hasil dari penelitian ini bisa dimanfaatkan dalam dunia kesehatan menjadi salah satu pertimbangan mengoptimalakan metode pengobatan kanker menggunakan kemoterapi. Khususnya pada kemoterapi berbasis senyawa *5-Fluorouracil*. Hasil penelitian juga diharapkan dapat menekan efek samping yang terjadi setelah melakukan kemoterapi. Dalam bidang keilmuan penelitian ini juga dapat dimanfaatkan dalam mnambah wawasan keilmuan statistika sebagai pengembangan dari penerapan metode regresi multivariat dan regresi logistik biner sebagai bahan referensi untuk penelitian-penelitian selanjutnya.

(Halaman ini sengaja dikosongkan)

BAB II TINJAUAN PUSTAKA

2.1 Analisis Regresi Multivariat

Model regresi multivariat adalah model regresi dengan lebih dari satu variabel respon yang saling berkorelasi dan satu atau lebih variabel prediktor (Johnson & Wichern, 2007). Misalkan terdapat variabel respon berjumlah q yaitu Y_1, Y_2, \dots, Y_q dan variabel prediktor sebanyak p yaitu X_1, X_2, \dots, X_p , maka model linear multivariat respon ke- q adalah:

$$\begin{aligned} Y_{1i} &= \beta_{01} + \beta_{11}X_1 + \dots + \beta_{p1}X_p + \varepsilon_{1i} \\ Y_{2i} &= \beta_{02} + \beta_{12}X_1 + \dots + \beta_{p2}X_p + \varepsilon_{2i} \\ &\vdots \\ Y_{qi} &= \beta_{0q} + \beta_{1q}X_1 + \dots + \beta_{pq}X_p + \varepsilon_{qi} \end{aligned} \quad (2.1)$$

dimana,

Y_{qi} = variabel respon ke- q

X_p = variabel prediktor ke- p

$\beta_{0q}, \beta_{1q}, \dots, \beta_{pq}$ = persamaan regresi yang nilainya belum diketahui

ε_{qi} = error persamaan, $i = 1, 2, \dots, n$

Model regresi multivariat yang terdiri atas q model linear dapat dituliskan dalam bentuk matriks seperti persamaan (2.2) sebagai berikut.

$$Y_{(n \times q)} = X_{n \times (p-1)} \cdot \beta_{(p-1) \times q} + \varepsilon_{n \times q} \quad (2.2)$$

Dengan,

$E(\varepsilon_1) = 0$ dan

$cov(\varepsilon_{(i)}, \varepsilon_{(j)}) = \sigma_{ii} I$

dimana,

n = banyaknya observasi

p = banyaknya variabel prediktor

q = banyaknya variabel respon

2.1.1 Uji Independensi

Variabel Y_1, Y_2, \dots, Y_q dikatakan bersifat saling bebas (independen) jika matriks korelasi antar variabel membentuk matriks identitas. Untuk menguji kebebasan antar variabel ini dapat dilakukan uji *Bartlett Sphericity* berikut. Uji ini berfungsi untuk mengetahui besarnya nilai korelasi antar variabel, dengan hipotesis sebagai berikut.

$H_0 : \rho = I$ (matriks korelasi sama dengan matriks identitas atau data independen)

$H_1 : \rho \neq I$ (matriks korelasi tidak sama dengan matriks identitas atau data dependen)

Statistik Uji :

$$\chi^2 = -\left\{n-1 - \frac{2p+5}{6}\right\} \ln |R| \quad (2.3)$$

Dimana,

n = jumlah observasi

p = jumlah variabel

$|R|$ = determinan dari matrik korelasi

Daerah kritis : Tolak H_0 jika, $\chi^2 > \chi^2_{(\alpha, \frac{1}{2}q(q-1))}$ sehingga dapat

disimpulkan bahwa variabel bersifat dependen. Berikut adalah rumus untuk menghitung nilai korelasinya

$$r_{ik} = \frac{S_{ik}}{\sqrt{S_{ii}} \sqrt{S_{kk}}} = \frac{\sum_{j=1}^n (x_{ij} - \bar{x}_i)(x_{jk} - \bar{x}_k)}{\sqrt{\sum_{j=1}^n (x_{ij} - \bar{x}_i)^2} \sqrt{\sum_{j=1}^n (x_{jk} - \bar{x}_k)^2}}; -1 \leq r_{ik} \leq 1 \quad (2.4)$$

Untuk

$i=1,2,\dots,p, \quad k=1,2,3,\dots,p, \quad r_{ik} = r_{ki}$ untuk semua i dan k .

2.1.2 Uji Distribusi Normal Multivariat

Asumsi lain yang harus dipenuhi dalam pemodelan regresi multivariat adalah variabel respon yang memiliki distribusi multivariat normal. Pemeriksaan distribusi multivariat normal

dilakukan untuk mengetahui apakah distribusi sebuah data mengikuti dan mendekati distribusi normal multivariat. (Johnson & Wichern, 2007). Terdapat beberapa teknik atau metode pemeriksaan asumsi distribusi normal multivariat Pemeriksaan distribusi multivariat normal dapat dilakukan dengan cara melihat korelasi yang diperoleh dari nilai koefisien korelasi plot *chi-square*. Yang kemudian dilihat kelurusan titik-titik yang membentuk garis pada q-q plot. Berikut ini adalah langkah-langkah untuk menggambar plot *chi-square*.

1. Menghitung nilai square distance (d_i^2). Nilai distance dapat dihitung dengan persamaan sebagai berikut.

$$d_i^2 = (y_i - \bar{y})^T S^{-1} (y_i - \bar{y}) ; i = 1, 2, \dots, n \quad (2.5)$$

2. Mengurutkan nilai square distance dari yang terkecil sampai yang terbesar.
3. Menentukan nilai

$$q_{c,p} \left(\left(i - \frac{1}{2} \right) / n \right) = \chi_p^2 \left(\left(n - i + \frac{1}{2} \right) / n \right) \quad (2.6)$$

4. Membuat plot antara nilai square distance (d_i^2)

$$\text{dengan } \chi_p^2 \left(\left(n - i + \frac{1}{2} \right) / n \right)$$

5. Melihat nilai t (proporsi nilai $d_i^2 < \chi_p^2 \left(\left(n - i + \frac{1}{2} \right) / n \right)$

terhadap jumlah pengamatan), jika nilai $t = d_i^2 <$

$$\chi_p^2 \left(\left(n - i + \frac{1}{2} \right) / n \right) \text{ atau } t \approx 50\% \text{ maka data tersebut sudah}$$

mengikuti distribusi normal multivariat.

Plot tersebut akan membentuk garis lurus jika data berdistribusi normal multivariat dan jika terdapat lengkungan menunjukkan penyimpangan dari normalitas.

2.1.3 Estimasi Parameter

Model regresi multivariat yang terdiri dari q model linear secara simultan dapat ditunjukkan bentuk matriks pada persamaan $Y_{(n \times p)} = X_{n \times (p-1)} \cdot \beta_{(p-1) \times q} + \varepsilon_{(n \times q)}$, β merupakan suatu matriks parameter regresi dengan ukuran $(p + 1) \times q$ dengan estimasinya adalah sebagai berikut

$$\beta = (X^T X)^{-1} X^T Y \quad (2.7)$$

Sedangkan ε merupakan matriks residual ditentukan oleh estimasi $\hat{\varepsilon} = Y - X\hat{\beta}$ (Johnson dan Wichern, 2007). Berikut ini notasi matriks untuk model regresi multivariat.

$$Y = \begin{matrix} n \times q \\ \begin{pmatrix} y_{11} & y_{12} & \cdots & y_{1q} \\ y_{21} & y_{22} & \cdots & y_{2q} \\ \vdots & \vdots & \ddots & \vdots \\ y_{n1} & y_{n2} & \cdots & y_{nq} \end{pmatrix} \end{matrix} = \begin{pmatrix} y_1' \\ y_2' \\ \vdots \\ y_q' \end{pmatrix}$$

$$X = \begin{matrix} (n \times (p+1)) \\ \begin{pmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1p} \\ 1 & x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{np} \end{pmatrix} \end{matrix}$$

$$\beta = \begin{matrix} ((p+1) \times q) \\ \begin{pmatrix} \beta_{01} & \beta_{02} & \cdots & \beta_{0q} \\ \beta_{11} & \beta_{12} & \cdots & \beta_{1q} \\ \vdots & \vdots & \ddots & \vdots \\ \beta_{p1} & \beta_{p2} & \cdots & \beta_{pq} \end{pmatrix} \end{matrix}$$

$$\varepsilon = \begin{matrix} (n \times q) \\ \begin{pmatrix} \varepsilon_{11} & \varepsilon_{12} & \cdots & \varepsilon_{1q} \\ \varepsilon_{21} & \varepsilon_{22} & \cdots & \varepsilon_{2q} \\ \vdots & \vdots & \ddots & \vdots \\ \varepsilon_{n1} & \varepsilon_{n2} & \cdots & \varepsilon_{nq} \end{pmatrix} \end{matrix}$$

Pada model regresi multivariat melakukan penaksiran parameter β menggunakan estimasi kuadrat terkecil. Penaksir estimasi parameter β yang dilakukan menggunakan prinsip meminimumkan jumlah kuadrat eror (Rencher, 2002).

2.1.4 Pengujian Signifikansi Parameter Model

Pengujian Signifikansi Parameter Model Terdapat dua uji yang dilakukan untuk mengetahui tingkat signifikansi yaitu parameter regresi signifikan terhadap model secara serentak dan secara parsial. Pengujian signifikan parameter dilakukan serentak dan parsial secara multivariat dengan menggunakan uji Wilk's Lambda.

a. Pengujian Serentak

Pengujian signifikan serentak dilakukan secara multivariat dengan menggunakan uji Wilk's Lambda. Pengujian signifikan serentak secara multivariat dilakukan bertujuan untuk mengetahui apakah semua estimasi parameter tidak sama dengan nol atau signifikan secara keseluruhan dalam model. Ukuran yang dinyatakan dalam regresi multivariat adalah Wilk's Lambda dan hipotesisnya adalah sebagai berikut (Rencher, 2002).

Hipotesis yang digunakan adalah sebagai berikut:

$$H_0 : \beta_1 = \beta_2 = \dots = \beta_j = \dots \beta_p = 0, \quad j = 1, 2, \dots, p$$

$$H_1 : \text{paling sedikit ada satu } \beta_j \neq 0$$

Hipotesis akan diuji menggunakan Statistik uji seperti persamaan (2.8) sebagai berikut

$$\Lambda = \frac{|E|}{|E + H|} \quad (2.8)$$

Dimana,

$$\begin{aligned} E &= Y^T Y - Y^T X (X^T X)^{-1} X^T Y \\ H &= (Y - X_i \beta_{(i)})^T (Y - X_i \beta_{(i)}) - (Y - X_i \beta)^T (Y - X_i \beta) \end{aligned} \quad (2.9)$$

Λ adalah Wilk's Lambda

\bar{y} adalah vektor rata-rata Y .

Keputusan : H_0 dapat ditolak jika $\Lambda_{hitung} \leq \Lambda_{\alpha,q,p,n-p-1}$ dan dapat disimpulkan bahwa parameter tidak sama dengan nol sehingga model signifikan. Nilai $\Lambda_{\alpha,q,p,n-p-1}$ merupakan nilai tabel kritis untuk Wilk's Lambda.

b. Pengujian Parsial

Pengujian signifikan parameter secara parsial ini bertujuan untuk melihat pengaruh signifikan setiap variabel prediktor terhadap variabel respon secara parsial. Ukuran yang dinyatakan dalam regresi multivariat adalah Wilk's Lambda dan hipotesisnya adalah sebagai berikut (Rencher, 2002).

Hipotesis:

$$H_0 : \beta_j = 0, j = 1, 2, \dots, p$$

$$H_1 : \beta_j \neq 0$$

Hipotesis akan diuji menggunakan Statistik uji seperti persamaan (2.15) sebagai berikut

$$\Lambda = \frac{|E|}{|E + H|} = \frac{\left| Y^T Y - \hat{\beta}^T X^T Y \right|}{\left| Y^T Y - \hat{\beta}_p^T X_p^T Y \right|} \quad (2.10)$$

Λ adalah Wilk's Lambda

\bar{y} adalah vektor rata-rata Y .

Keputusan : H_0 dapat ditolak jika $\Lambda \leq \Lambda_{\alpha,q,p,n-p-1}$ dan dapat disimpulkan bahwa parameter tidak sama dengan nol sehingga model signifikan. Nilai $\Lambda_{\alpha,q,p,n-p-1}$ merupakan nilai tabel kritis untuk Wilk's Lambda.

2.1.5 Uji Asumsi Residual IIDN

Setelah melakukan uji signifikansi, residual pada model regresi perlu dilakukan pemeriksaan asumsi Identik, Independen dan mengikuti distribusi Normal (IIDN)

a. Uji Asumsi Residual Identik

Asumsi yang harus dipenuhi dalam melakukan pemodelan regresi multivariat adalah residual memiliki matriks varian-kovarian yang homogen. Untuk menguji asumsi identik dalam regresi multivariat ini dapat dipergunakan statistik uji Box's M (Rencher, 2002)

Hipotesis yang digunakan adalah sebagai berikut:

$$H_0 : \Sigma_1 = \Sigma_2 = \dots = \Sigma_k$$

$$H_1 : \text{Minimal ada satu } \Sigma_k \neq \Sigma_l \text{ untuk } k \neq l$$

$$\text{Statistik Uji : } \chi^2_{hitung} = u = -2(1 - c_1) \ln M \quad (2.11)$$

Dimana,

$$\ln M = \frac{1}{2} \sum_{i=1}^k v_i \ln |S_i| - \frac{1}{2} \left(\sum_{i=1}^k v_i \right) \ln |S_{pool}| \quad (2.12)$$

$$S_{pool} = \frac{\sum_{i=1}^k v_i S_i}{\sum_{i=1}^k v_i} \quad (2.13)$$

$$c_1 = \left[\sum_{i=1}^k \frac{1}{v_i} - \frac{1}{\sum_{i=1}^k v_i} \right] \left[\frac{2p^2 + 3p - 1}{6(p+1)(k-1)} \right] \quad (2.14)$$

$$v_i = n_i - 1 \quad (2.15)$$

Keterangan :

z = banyak variabel residual.

k = banyaknya kelompok.

n_i = banyak pengamatan pada kelompok ke- i .

S_i = matrik varian kovarians dari kelompok ke- i .

Daerah Penolakan : Gagal tolak H_0 jika $u \leq \chi^2_{\alpha; \frac{1}{2}(k-1)z(z+1)}$ atau

dapat disimpulkan matriks varian kovarian residual bersifat homogen.

b. Uji Asumsi Residual Independen

Tujuan dilakukan pengujian residual independen adalah untuk mengetahui apakah korelasi antar residual bernilai 0 atau tidak. Untuk menguji korelasi antar residual dapat dilakukan uji Bartlett Sphericity (Maarrison, 2005). Uji ini berfungsi untuk mengetahui besarnya nilai korelasi antar variabel, dengan hipotesis sebagai berikut.

$H_0 : \rho = I$ (matriks korelasi sama dengan matriks identitas atau data depeneden)

$H_1 : \rho \neq I$ (matriks korelasi tidak sama dengan matriks identitas atau data independen)

Statistik Uji :

$$\bar{r}_k = \frac{1}{p-1} \sum_{i=1}^p r_{ik}, k = 1, 2, \dots, p \quad (2.16)$$

$$\bar{r} = \frac{2}{p(p-1)} \sum_{i < k} r_{ik} \quad (2.17)$$

$$\hat{\gamma} = \frac{(p-1)^2 [1 - (1 - \bar{r})^2]}{p - (p-2)(1 - \bar{r})^2} \quad (2.18)$$

Dimana :

\bar{r}_k = rata-rata elemen diagonal pada kolom atau baris ke k dari matrik R (matrik korelasi)

\bar{r} = rata-rata keseluruhan dari elemen diagonal

Daerah kritis : Tolak H_0 jika,

$$T = \frac{(n-1)}{(1-\bar{r})^2} \left[\sum_{i < k} (r_{ik} - \bar{r})^2 - \hat{\gamma} \sum_{k=1}^p (\bar{r}_k - \bar{r})^2 \right] > \chi^2_{(p+1)(p-2)/2; \alpha} \quad (2.19)$$

Keputusan: Tolak H_0 jika variabel-variabel saling berkorelasi hal ini berarti terdapat hubungan antar variabel.(Morrison, 2005).

c. Uji Asumsi Residual Berdistribusi Normal Multivariat

Pemeriksaan distribusi multivariat normal dapat dilakukan dengan cara menghitung jarak kuadrat untuk setiap pengamatan (Johnson & Wichern, 2007). Yang kemudian dilihat kelurusan titik-titik yang membentuk garis pada q-q plot. Berikut ini adalah langkah-langkah untuk menggambar plot *chi-square*.

1. Menghitung nilai square distance (d_i^2). Nilai distance dapat dihitung dengan persamaan sebagai berikut.

$$d_i^2 = (\hat{\varepsilon}_i - \bar{\varepsilon})^T S^{-1} (\hat{\varepsilon}_i - \bar{\varepsilon}) ; i = 1, 2, \dots, n$$

2. Mengurutkan nilai square distance dari yang terkecil sampai yang terbesar.

3. Menentukan nilai $q_{c,p} \left(\left(i - \frac{1}{2} \right) / n \right) = \chi_p^2 \left(\left(n - i + \frac{1}{2} \right) / n \right)$

4. Membuat plot antara nilai square distance (d_i^2) dengan

$$\chi_p^2 \left(\left(n - i + \frac{1}{2} \right) / n \right)$$

5. Melihat nilai t (proporsi nilai $d_i^2 < \chi_p^2 \left(\left(n - i + \frac{1}{2} \right) / n \right)$)

terhadap jumlah pengamatan), jika nilai $t = d_i^2 <$

$$\chi_p^2 \left(\left(n - i + \frac{1}{2} \right) / n \right) \text{ atau } t \approx 50\% \text{ maka data tersebut sudah}$$

mengikuti distribusi normal multivariat.

Plot tersebut akan membentuk garis lurus jika data berdistribusi normal multivariat dan jika terdapat lengkungan menunjukkan penyimpangan dari normalitas.

2.1.6 Ukuran Kebaikan Model Regresi Multivariat

Pada analisis regresi multivariat, ukuran yang digunakan untuk mengukur kebaikan model yang menjelaskan proporsi variabilitas variabel respon dengan menggunakan pendekatan

ukuran Manova One-Way yaitu Eta Square Lambda. Berikut ini persamaan untuk menghitung nilai Eta Square Lambda.

$$\eta_{\Lambda}^2 = 1 - \Lambda \quad (2.20)$$

Dengan Λ adalah nilai wilk's lambda dan η_{Λ}^2 adalah nilai keterkaitan antar variabel respon dan variabel prediktor dengan $0 < \eta_{\Lambda}^2 < 1$ Artinya apabila nilai η_{Λ}^2 semakin mendekati 1 berarti ukuran kebaikan model yang dijelaskan oleh proporsi variabilitas variabel respon semakin besar (Rencher, 2002).

2.2 Regresi Linear Berganda

Regresi linear berganda merupakan metode yang memodelkan hubungan antara variabel respon (y) dan variabel prediktor ($x_1, x_2, x_3, \dots, x_p$). Model regresi linear untuk p variabel prediktor secara umum ditulis sebagai berikut (Fotheringham, Brundson, & Charlton, 2002).

$$y_i = \beta_0 + \sum_{k=1}^p \beta_k X_{ik} + \varepsilon_i, \quad (2.21)$$

Dimana :

$i = 1, 2, 3, \dots, n$

y_i = nilai observasi variabel respon pada pengamatan ke- i

X_{ik} = nilai observasi variabel prediktor ke- k pada pengamatan ke- i , dengan $k = 1, 2, \dots, p$

β_0 = nilai intersep model regresi

β_k = koefisien regresi variabel prediktor ke- k

ε_i = *error* pada pengamatan ke- i dengan asumsi independen, identik, dan berdistribusi normal, dengan *mean* nol dan varians konstan

Pendugaan parameter model regresi linear diperoleh dengan menggunakan metode *Ordinary Least Square* (OLS). Metode OLS didapat dengan meminimumkan jumlah kuadrat *error*. Pendugaan parameter model didapat dari persamaan sebagai berikut (Drapper & Smith, 1991).

$$\hat{\beta} = (X^T X)^{-1} X^T Y, \quad (2.22)$$

dengan

$\hat{\beta}$ = vektor dari parameter yang diestimasi berukuran
 $(p + 1) \times 1$

X = matriks variabel prediktor berukuran $n \times (p + 1)$

Y = vektor observasi dari variabel respon berukuran $n \times 1$

2.2.1 Pengujian Hipotesis Model Regresi Linear Berganda

Pengujian parameter dilakukan untuk mengetahui apakah suatu variabel memberikan pengaruh secara signifikan terhadap model. Pengujian parameter dapat dilakukan secara serentak dan individu (parsial) dengan penjelasan sebagai berikut.

a. Uji Serentak

Uji serentak adalah uji signifikansi model secara keseluruhan atau untuk mengetahui apakah semua variabel prediktor yang dimasukkan ke dalam model memberikan pengaruh secara bersama-sama. Uji serentak dapat dilakukan menggunakan tabel Anova seperti terlihat pada Tabel 2.1.

Tabel 2.2 Analisis ragam (ANOVA) Uji Parameter

Sumber variasi	Df	Sum of Square (SS)	Mean Square (MS)	F_{hitung}
Regresi	p	$\sum_{i=1}^n (\hat{y}_i - \bar{y})^2$	$\frac{SS_{regresi}}{df_{regresi}}$	
Error	$n - p - 1$	$\sum_{i=1}^n (y_i - \hat{y}_i)^2$	$\frac{SS_{error}}{df_{error}}$	$\frac{MS_{regresi}}{MS_{error}}$
Total	$n - 1$	$\sum_{i=1}^n (y_i - \bar{y})^2$	–	

Hipotesis yang digunakan untuk pengujian secara serentak adalah sebagai berikut.

$$H_0 : \beta_1 = \beta_2 = \dots = \beta_p = 0$$

$$H_1 : \text{minimal terdapat satu } \beta_k \neq 0, k = 1, 2, \dots, p$$

Statistik uji yang digunakan sebagai berikut.

$$F_{hitung} = \frac{MS_{regresi}}{MS_{residual}} \quad (2.23)$$

Daerah penolakan, tolak H_0 jika $F_{hitung} > F_{\alpha; p, (n-p-1)}$ (Drapper & Smith, 1991).

b. Uji Parsial

Uji parsial adalah uji signifikansi masing-masing variabel prediktor secara individual. Hipotesis yang digunakan dalam uji individu adalah sebagai berikut.

$$H_0 : \beta_k = 0$$

$$H_1 : \beta_k \neq 0, k = 1, 2, \dots, p$$

Statistik uji yang digunakan sebagai berikut.

$$t_{hitung} = \frac{\hat{\beta}_k}{SE(\hat{\beta}_k)} \quad (2.24)$$

dimana $SE(\hat{\beta}_k)$ adalah standar error dari $\hat{\beta}_k$. Daerah penolakan yaitu tolak H_0 jika $|t_{hitung}| > t_{\frac{\alpha}{2}; (n-p-1)}$ atau tolak H_0 jika

$p_{value} < \alpha$ yang berarti bahwa parameter berpengaruh secara signifikan terhadap model (Drapper & Smith, 1991).

2.2.2 Pengujian Asumsi Residual

Pengujian asumsi residual (*Goodness of fit*) dilakukan untuk mengetahui apakah residual yang dihasilkan telah memenuhi asumsi yakni identik, independen, dan berdistribusi normal (IIDN). Asumsi identik terpenuhi jika varians antar residual homogen yakni σ^2 dan tidak terjadi heteroskedastisitas (Gujarati, 2003). Asumsi klasik kedua yang harus dipenuhi adalah tidak terdapat korelasi antar residual yang ditunjukkan oleh nilai kovarian antara ϵ_i dan ϵ_j sama dengan nol. Residual dari model regresi harus mengikuti distribusi normal dengan *mean* nol dan varians σ^2 . Uji asumsi distribusi normal dapat dilakukan menggunakan uji *Kolmogorov-smirnov*.

a. Asumsi Identik

Asumsi identik (homoskedastisitas) berarti bahwa varians pada residual sama atau identik. Kebalikannya adalah kasus heteroskedastisitas, yaitu jika kondisi varians *residual* tidak identik (Gujarati, 2003).

$$\text{var}(y_i) = \text{var}(\varepsilon_i) = \sigma^2 ; i = 1, 2, \dots, n \quad (2.25)$$

Uji identik dapat menggunakan uji Glejser. Hipotesis yang digunakan adalah sebagai berikut:

$H_0 : \sigma_1^2 = \sigma_2^2 = \dots = \sigma_n^2 = \sigma^2$ (residual identik)

$H_1 : \text{Minimal ada satu } \sigma_i^2 \neq \sigma ; i = 1, 2, \dots, n$ (residual tidak identik)

Statistik uji yang digunakan adalah seperti berikut,

$$F_{hitung} = \frac{\frac{\sum_{i=1}^n (|\varepsilon_i| - |\bar{\varepsilon}|)^2}{v-1}}{\frac{\sum_{i=1}^n (|\varepsilon_i| - |\hat{\varepsilon}_i|)^2}{n-v}} \quad (2.26)$$

dimana nilai v adalah banyaknya parameter model Glejser dan nilai $v = p + r$. Jika $F_{hitung} > F_{\alpha; (v-1, n-v)}$ atau $p\text{-value} < \alpha$ maka tolak H_0 yang berarti bahwa tidak terindikasi terdapat kasus homoskedastisitas dan sebaliknya jika $F_{hitung} < F_{\alpha; (v-1, n-v)}$ atau $p\text{-value} > \alpha$ maka gagal tolak H_0 yang berarti bahwa terindikasi terdapat kasus homoskedastisitas.

b. Asumsi Independen

Asumsi independen merupakan asumsi dari model regresi yang mengharuskan tidak terdapat korelasi antar residual. Uji yang digunakan untuk mendeteksi kasus autokorelasi adalah uji *Durbin-Watson* (Drapper & Smith, 1991). Hipotesis yang digunakan adalah sebagai berikut.

Hipotesis:

$H_0 : \rho = 0$ (tidak terjadi autokorelasi)

$H_1 : \rho \neq 0$ (terjadi autokorelasi)

Statistik uji yang digunakan:

$$d = \frac{\sum_{i=2}^n (e_i - e_{i-1})^2}{\sum_{i=1}^n e_i^2} \quad (2.27)$$

Langkah selanjutnya adalah membandingkan hasil statistic uji dengan tabel Durbin Watson (DW). Tabel DW terdiri atas batas bawah (d_L) dan batas atas (d_U). Berikut beberapa keputusan setelah membandingkan dengan tabel DW.

$$\begin{aligned} d < d_L & : \text{tolak } H_0 \\ d > 4 - d_L & : \text{tolak } H_0 \\ d_U < d < 4 - d_L & : \text{terima } H_0 \\ d_L \leq d \leq d_U \text{ dan } 4 - d_U \leq d \leq 4 - d_L & \end{aligned} \quad (2.28)$$

c. Uji Normalitas *Kolmogorov-Smirnov*

Uji *Kolmogorov-Smirnov* bertujuan untuk mengetahui apakah suatu data telah mengikuti suatu distribusi tertentu (Drapper & Smith, 1991). Hipotesis yang digunakan

$$H_0 : F_n(\varepsilon) = F_0(\varepsilon)$$

$$H_1 : F_n(\varepsilon) \neq F_0(\varepsilon)$$

atau

H_0 : residual berdistribusi normal

H_1 : residual tidak berdistribusi normal

Statistik uji:

$$D = \sup_{\varepsilon} |F_n(\varepsilon) - F_0(\varepsilon)| \quad (2.29)$$

Tolak H_0 apabila $D > D_\alpha$. D_α adalah nilai kritis untuk uji *Kolmogorov Smirnov* satu sampel, diperoleh dari tabel *Kolmogorov Smirnov* satu sampel, $F_n(\varepsilon)$ adalah nilai peluang kumulatif (fungsi distribusi kumulatif) berdasarkan data sampel, $F_0(\varepsilon)$ adalah nilai peluang kumulatif (fungsi distribusi kumulatif) dibawah H_0 .

2.2.3 Ukuran Kebaikan Model Regresi Linear

Pada analisis regresi linear, ukuran yang digunakan untuk mengukur kebaikan model yang menjelaskan proporsi variabilitas variabel respon dengan menggunakan nilai R^2 . Berikut ini adalah rumus untuk mendapatkan nilai R^2 .

$$R^2 = \frac{SSR}{SSE} = \frac{\sum_{i=1}^n (y_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (2.30)$$

R^2 adalah ukuran kebaikan model antar variabel respon dan variabel prediktor dengan $0 \leq R^2 \leq 1$ Artinya apabila nilai R^2 semakin mendekati 1 berarti ukuran kebaikan model yang dijelaskan oleh proporsi variabilitas variabel respon semakin besar.

2.3 Regresi Logistik Biner

Regresi logistik biner adalah salah satu metode statistik yang biasanya digunakan untuk menjelaskan hubungan antara variabel respon (y) yang bersifat biner atau dikotomus dengan satu atau lebih variabel prediktor (x) yang bersifat metrik atau non-metrik (Hosmer & Lomeshoe, 2000). Regresi logistik biner hanya digunakan untuk kasus khusus yaitu apabila variabel respon (Y) merupakan variabel kualitatif yang bersifat biner atau dikotomus. Variabel dikotomus adalah variabel yang hanya mempunyai dua kemungkinan nilai, misalnya sukses dan gagal.

Untuk memudahkan, maka variabel respon/dependen diberi notasi Y dan variabel prediktor/independen dinotasikan X . Apabila Y menghasilkan dua kategori, misalnya “1” jika sukses dan “0” jika gagal. Variabel respon Y berdistribusi Binomial dengan parameter π_i , dimana untuk setiap pengamatan ke- i ditulis pada persamaan (2.31) (Agresti, 2007).

$$y_i \sim \text{Binomial}(1, \pi_i) \quad (2.31)$$

dengan fungsi probabilitas yang ditunjukkan pada persamaan (2.32).

$$f(y_i) = (\pi_i(\mathbf{x}_i))^{y_i} (1 - \pi_i(\mathbf{x}_i))^{1-y_i}, y_i = 0, 1 \quad (2.32)$$

dimana $\pi_i(\mathbf{x}_i)$ adalah probabilitas dari kejadian ke- i . Apabila $y_i=1$, maka $f(y_i) = \pi_i(\mathbf{x}_i)$

dan apabila $y_i = 0$, maka $f(y_i) = 1 - \pi_i(\mathbf{x}_i)$.

Dalam regresi logistik, hubungan antara variabel prediktor dan variabel respon bukanlah suatu fungsi linear. Bentuk persamaan regresi logistik ditunjukkan pada persamaan (2.33).

$$\pi_i(\mathbf{x}_i) = \frac{\exp(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + K + \beta_p x_{ip})}{1 + \exp(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + K + \beta_p x_{ip})} \quad (2.33)$$

$$\pi_i(\mathbf{x}_i) = \frac{\exp(\mathbf{x}_i^T \boldsymbol{\beta})}{1 + \exp(\mathbf{x}_i^T \boldsymbol{\beta})}$$

Dimana

β_0 = konstanta,

β_j = koefisien regresi

j = banyaknya variabel prediktor.

Terdapat suatu bentuk alternatif dari persamaan regresi logistik seperti persamaan (2.34) yang merupakan transformasi logit dari $\pi(\mathbf{x}_i)$ (Yan & Su, 2009)

$$\text{Logit}[\pi_i(\mathbf{x}_i)] = (\mathbf{x}_i^T \boldsymbol{\beta}) = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + K + \beta_p x_{ip} \quad (2.34)$$

2.3.1 Estimasi Parameter

Estimasi parameter dalam regresi logistik dilakukan dengan metode *Maximum Likelihood* (MLE). Metode MLE memberikan nilai estimasi β dengan memaksimumkan fungsi likelihood (Hosmer et al., 2013). Metode tersebut mengestimasi parameter β dengan cara memaksimumkan fungsi likelihood dan mensyaratkan bahwa data harus mengikuti suatu distribusi tertentu. Pada regresi logistik, setiap pengamatan mengikuti distribusi bernoulli sehingga dapat ditentukan fungsi likelihoodnya.

$$f(x_i) = \pi(x_i)^{y_i} (1 - \pi(x_i))^{1-y_i} \quad (2.35)$$

Metode *MLE* digunakan karena distribusi dari variabel respon telah diketahui yaitu distribusi Bernoulli. *MLE* didapatkan dengan cara memaksimumkan logaritma fungsi likelihood.

$$\text{Logit}[\pi_i(\mathbf{x}_i)] = (\mathbf{x}_i^T \boldsymbol{\beta}) = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_p x_{ip} \quad (2.36)$$

Dari persamaan (2.36) didapatkan fungsi likelihood pada Persamaan (2.37) (Agresti, 2007).

$$L(\mathbf{X}, \boldsymbol{\beta}) = \prod_{i=1}^n [\pi_i(\mathbf{x}_i)]^{y_i} [1 - \pi_i(\mathbf{x}_i)]^{1-y_i} \quad (2.37)$$

$$\ln L(\mathbf{X}, \boldsymbol{\beta}) = \sum_{i=1}^n \left[y_i (\mathbf{x}_i \boldsymbol{\beta}) - \ln(1 + \exp(\mathbf{x}_i \boldsymbol{\beta})) \right]$$

Hasil estimasi parameter $\hat{\boldsymbol{\beta}}$ adalah sebagai berikut.

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}^T \mathbf{W} \mathbf{X})^{-1} \mathbf{X}^T \mathbf{W} \mathbf{z} \quad (2.38)$$

dimana \mathbf{z} merupakan vektor $n \times 1$ dengan :

$$z_i = \text{Logit}[\hat{\pi}_i(\mathbf{x}_i)] + \frac{y_i - \hat{\pi}_i(\mathbf{x}_i)}{\hat{\pi}_i(\mathbf{x}_i)[1 - \hat{\pi}_i(\mathbf{x}_i)]} \quad (2.39)$$

Matriks kovarian untuk $\hat{\beta}$ ditampilkan pada persamaan sebagai berikut.

$$\text{Var}(\hat{\beta}) = (\mathbf{X}^T \text{diag} [\hat{\pi}_i (1 - \hat{\pi}_i)] \mathbf{X})^{-1} \quad (2.40)$$

2.3.2 Pengujian Signifikansi Parameter

Pengujian signifikansi parameter digunakan untuk mengetahui variabel prediktor yang berpengaruh terhadap variabel respon. Pengujian ini dilakukan dengan cara uji serentak (bersama-sama) dan uji parsial (sendiri-sendiri). Pengujian signifikan serentak dilakukan secara menggunakan *Likelihood Ratio Test* sedangkan untuk pengujian secara parsial dengan menggunakan metode *Wald Test*.

a. Uji Serentak

Uji serentak bertujuan untuk mengetahui apakah model signifikan dan memeriksa pengaruh variabel prediktor di dalam model secara bersama-sama. Uji serentak dapat dilakukan menggunakan uji Chi-Square. Berikut ini adalah hipotesis uji serentak.

$$H_0: \beta_1 = \beta_2 = \dots = \beta_i = 0$$

$$H_1: \text{Paling tidak terdapat satu } \beta_i \neq 0; i = 1, 2, \dots, p$$

Statistik uji:

$$G = -2 \ln \frac{\left(\frac{n_1}{n} \right)^{n_1} \left(\frac{n_0}{n} \right)^{n_0}}{\sum_{i=1}^n \hat{\pi}_i^{y_i} (1 - \hat{\pi}_i)^{(1-y_i)}} \quad (2.41)$$

Dimana,

$$n_1 = \sum_{i=1}^n y_i \quad n_0 = \sum_{i=1}^n (1 - y_i) \quad n = n_1 + n_0$$

Statistik uji G adalah merupakan *Likelihood Ratio Test* dimana nilai G mengikuti distribusi *Chi-Squared* sehingga H_0

ditolak jika $G > \chi^2_{(v,\alpha)}$ dengan v derajat bebas adalah banyaknya parameter dalam model tanpa β_0 (Hosmer et al., 2013).

b. Uji Parsial

Setelah dilakukan pengujian secara serentak dari koefisien parameter β terhadap variabel respon, langkah analisis dilanjutkan dengan pengujian signifikansi variabel prediktor secara parsial terhadap variabel respon. Pengujian signifikansi secara parsial dilakukan dengan metode *Wald Test* untuk mengetahui variabel-variabel prediktor yang signifikan terhadap peluang sukses. Hipotesis yang digunakan untuk uji ini adalah

$H_0 : \beta_j = 0$ (variabel ke- j tidak berpengaruh signifikan)

$H_1 : \beta_j \neq 0, j = 1, 2, \dots, p$ (variabel ke- j berpengaruh signifikan)

Statistik uji :

$$W^2 = \frac{\hat{\beta}_i^2}{SE(\hat{\beta}_i)^2} \quad (2.42)$$

dengan $SE(\hat{\beta}_j) = \sqrt{\text{var}(\hat{\beta}_j)}$

Daerah kritis : Statistik uji tersebut mengikuti distribusi *Chi-Squared* sehingga H_0 ditolak jika $W^2 > \chi^2_{(v,\alpha)}$ dengan v *degrees of freedom* banyaknya prediktor atau nilai $p\text{-value} < \alpha$. Artinya, variabel ke- j berpengaruh signifikan terhadap pembentukan model (Hosmer et al., 2013).

2.3.3 Interpretasi Model

Intepretasi terhadap koefisien parameter ini dilakukan untuk menentukan kecenderungan/hubungan fungsional antara variabel prediktor dengan variabel respon serta menunjukkan besarnya perubahan nilai variabel respon untuk setiap perubahan satu unit variabel prediktor (Hosmer et al., 2013). Dalam hal ini digunakan besaran *Odds ratio* atau e^β dan dinyatakan dengan ψ . *Odds ratio* diartikan sebagai kecenderungan variabel respon memiliki suatu

nilai tertentu jika diberikan $x=1$ dan dibandingkan pada $x=0$. Keputusan tidak terdapat hubungan antara variabel prediktor dengan variabel respon diambil jika nilai *Odds ratio* (ψ) = 1. Jika nilai *Odds ratio* (ψ) < 1, maka antara variabel prediktor dan variabel respon terdapat hubungan negatif setiap kali perubahan nilai variabel bebas (x) dan jika *Odds ratio* (ψ) > 1 maka antara variabel prediktor dengan variabel respon terdapat hubungan positif setiap kali perubahan nilai variabel bebas (x).

2.3.4 Evaluasi dan Validasi Klasifikasi Model

Evaluasi prosedur klasifikasi adalah suatu evaluasi yang melihat peluang kesalahan klasifikasi yang dilakukan oleh suatu fungsi klasifikasi (Johnson and Wichern, 2007). Validasi adalah proses mengevaluasi akurasi prediksi dari sebuah model, validasi mengacu untuk mendapatkan prediksi dengan menggunakan model yang ada kemudian membandingkan hasil yang diperoleh dengan hasil yang diketahui (Gorunescu, 2011). Mengevaluasi akurasi dari model klasifikasi sangat penting, akurasi dari sebuah model mengindikasikan kemampuan model tersebut untuk memprediksi class target (Vercellis, 2009). Untuk mengevaluasi model digunakan metode *confusion matrix*, dan kurva ROC (*Receiver Operating Characteristic*).

Confusion matrix memberikan rincian klasifikasi, kelas yang diprediksi akan ditampilkan di bagian atas matrix dan kelas yang diobservasi ditampilkan di bagian kiri (Gorunescu, 2011). Evaluasi model *confussion matrix* menggunakan tabel seperti matrix pada Tabel 2.1 sebagai berikut.

Tabel 2.1 Tabel Klasifikasi		
Hasil Observasi	Hasil Prediksi	
	π_1	π_2
π_1	n_{11}	n_{12}
π_2	n_{21}	n_{22}

Keterangan :

n_{11} = Jumlah subyek dari π_1 tepat diklasifikasi sebagai π_1

n_{12} = Jumlah subyek dari π_1 salah diklasifikasi sebagai π_2

n_{21} = Jumlah subyek dari π_2 salah diklasifikasi sebagai π_1

n_{22} = Jumlah subyek dari π_2 tepat diklasifikasi sebagai π_2

Perhitungan akurasi klasifikasi dengan confusion matrix dapat digunakan untuk komposisi data yang *balance*. Dengan menggunakan confusion matrix, beberapa pengukuran Kinerja klasifikasi dapat dilakukan, yaitu akurasi, sensitifitas, spesifitas, dan presisi. Akurasi menunjukkan efektifitas *classifier* secara menyeluruh. Semakin besar nilai akurasi, maka kinerja *classifier* semakin baik. Sensifitas digunakan untuk mengukur efektifitas sebuah *classifier* untuk mengidentifikasi kelas positif, sedangkan spesififikasi mengukur efektifitas *classifier* dalam mengidentifikasi kelas.

Akurasi klasifikasi pada data *unbalance* dapat dihitung menggunakan nilai *Area Under ROC Curve* (AUC). Untuk mendapatkan nilai AUC dapat dihitung dengan rumus seperti persamaan (2.43) dengan I adalah jumlah kelas klasifikasi yang terbentuk (Bekkar, Djemaa & Alitouch, 2013).

$$AUC = \frac{1}{I} \sum_{i=1}^I R_i \quad (2.43)$$

$$\text{dimana, } R_i = \frac{n_{ii}}{\sum_{l=1}^I n_{il}}, i = 1, 2, \dots, I \quad (2.44)$$

2.4 Tinjauan Non Statistik

Tinjauan non statistik pada penelitian meliputi Pemahaman tentang Kanker, Kanker kolorektal, Kemoterapi, *5-Fluorouracil* (*5-FU*), *White Blood Cells* dan *Neutrophil* dan IMT.

2.4.1 Kanker

Menurut WHO, kanker adalah istilah umum untuk satu kelompok besar penyakit yang dapat mempengaruhi setiap bagian

dari tubuh. Istilah lain yang digunakan adalah tumor ganas dan neoplasma. Salah satu fitur mendefinisikan kanker adalah pertumbuhan sel-sel baru secara abnormal yang tumbuh melampaui batas normal, dan yang kemudian dapat menyerang bagian sebelah tubuh dan menyebar ke organ lain. Menurut National Cancer Institute(2009), kanker adalah suatu istilah untuk penyakit di mana sel-sel membelah secara abnormal tanpa kontrol dan dapat menyerang jaringan di sekitarnya. Proses ini disebut metastasis. Metastasis merupakan penyebab utama kematian akibat kanker (WHO, 2018).

Kanker adalah istilah umum yang dipakai untuk menunjukkan neoplasma ganas, dan ada banyak tumor atau neoplasma lain yang tidak bersifat kanker (Price, 2006). Neoplasma secara harfiah berarti pertumbuhan baru. Suatu neoplasma, sesuai definisi Wills, adalah massa abnormal jaringan yang pertumbuhannya berlebihan dan tidak terkoordinasikan dengan pertumbuhan jaringan normal serta terus demikian walaupun rangsangan yang memicu perubahan tersebut telah berhenti (Kumar et al., 2007). Istilah tumor kurang lebih merupakan sinonim dari istilah neoplasma. Semua istilah tumor diartikan secara sederhana sebagai pembengkakan atau Universitas Sumatera Utara gumpalan, dan kadang-kadang istilah tumor sejati dipakai untuk membedakan neoplasma dengan gumpalan lainnya. Neoplasma dapat dibedakan berdasarkan sifat-sifatnya; ada yang jinak, ada pula yang ganas (Price, 2006).

2.4.3 Kemoterapi

Kemoterapi adalah salah satu prosedur perawatan yang paling umum diberikan untuk kanker. Terapi ini mengandalkan kemampuan dari obat-obat khusus untuk menghancurkan sel-sel kanker yang menyerang tubuh. Obat tersebut bekerja dengan memperlambat maupun menghentikan pertumbuhan sel kanker. Kemoterapi adalah penggunaan antipreparat antineoplastik sebagai upaya untuk membunuh sel-sel tumor dengan mengganggu fungsi dan reproduksi selular. Kemoterapi merupakan salah satu modalitas pengobatan pada kanker secara sistemik yang sering

dipilih terutama untuk mengatasi kanker stadium lanjut, lokal maupun metastatis (Desen, 2008). Kemoterapi sangat penting dan bermanfaat karena bersifat sistemik membunuh sel-sel kanker dengan cara melalui infus (Otto, 2003). Tujuan dari kemoterapi adalah penyembuhan, pengontrolan, dan paliatif harus realistis, karena tujuan tersebut akan menetapkan medikasi yang digunakan dan keagresifan rencana pengobatan. Obat kemoterapi secara umum disebut sitostatika, efeknya adalah membunuh atau menghambat semua sel yang sedang aktif membelah diri (Smeltzer, 2002). Menurut (Gale, 2000) Obat kemoterapi dapat diberikan dengan melalui topikal, oral, intravena, intramuskular, subkutan, arteri, intratekal. Pemberian biasanya tergantung pada tipe obat dosis yang dibutuhkan, jenis, lokasi dan luasnya tumor yang diobati.

Efek samping dari kemoterapi meliputi anemia, trombositopenia, leucopenia, neutropenia, mual dan muntah, alopesia (rambut rontok), stomatitis, reaksi alergi, neurotoksik dan ekstrasvasi. Efek samping kemoterapi dapat terjadi akut dan jangka panjang. Efek samping akut dapat terjadi beberapa jam sampai beberapa minggu setelah pemberian kemoterapi, berupa mielosupresi, mual, muntah, alopesia, mukositis orointestinal, kelainan fungsi hati, alergi serta ulserasi lokal. Tingkat kerusakan organ akibat efek samping kemoterapi berbeda pada tiap individu tergantung berbagai faktor antara lain jenis dan dosis kemoterapi yang dipakai, jangka waktu pemberian, faktor individu seperti ras, status gizi, keadaan organ tempat detoksikasi, dan ekskresi obat tersebut.

2.4.4 5-Fluorouracil (5-FU)

5-Fluorouracil (5-FU) merupakan agen kemoterapi utama yang digunakan untuk terapi kanker kolon. *5-FU* adalah antimetabolit yang bekerja secara antagonis dengan timin terhadap aktivitas enzim timidilat sintetase (TS). *5-FU* merupakan prodrug, metabolisme *5-FU* menghasilkan fluoridin-5-trifosfat (FUTP) yang bergabung ke dalam RNA dan mempengaruhi fungsinya, dan fluorodeoksiuridilat (FdUMP) yang menghambat replikasi DNA

(Longley & Johnston, 2007). Efek samping dari *5-FU* yang ditemukan pada pasien antara lain neutropenia, stomatitis, diare, dan *hand-food syndrome*. Masing-masing efek ini terkait dengan metode pemberian yang diterapkan pada pasien (Meyerhardt & Mayer, 2005). Pada kasus yang efek samping *5-FU* yang paling parah adalah kardi toksisitas meskipun hal ini jarang ditemui (Thomas *et al.*, 2004). Dibandingkan dengan agen kemoterapi yang lain, *5-FU* memiliki selektivitas yang tinggi pada aktivitas TS dan efek samping yang ditimbulkan relatif lebih ringan. Meskipun demikian, efektivitas *5-FU* sebagai agen kemoterapi baru mencapai 15% sehingga diperlukan pengembangan agen kokemoterapi untuk meningkatkan efektivitas terapi dengan *5-FU* (Meyerhardt & Mayer, 2005).

2.4.5 White Blood Cells (WBC)

White Blood Cells biasa disebut Leukosit atau *White Corpuscle*, merupakan sel darah putih yang diproduksi oleh jaringan hemopoetik untuk jenis bergranula (polimorfonuklear) dan jaringan limpatik untuk jenis tak bergranula (mononuklear), berfungsi dalam sistem pertahanan tubuh terhadap infeksi (Sutedjo, 2006). Leukosit paling sedikit dalam tubuh jumlahnya sekitar 4.000-11.000/mm³. Berfungsi untuk melindungi tubuh dari infeksi. Karena itu, jumlah leukosit tersebut berubah-ubah dari waktu ke waktu, sesuai dengan jumlah benda asing yang dihadapi dalam batas-batas yang masih dapat ditoleransi tubuh tanpa menimbulkan gangguan fungsi (Sadikin, 2002). Meskipun leukosit merupakan sel darah, tapi fungsi leukosit lebih banyak dilakukan di dalam jaringan. Leukosit hanya bersifat sementara mengikuti aliran darah ke seluruh tubuh. Apabila terjadi peradangan pada jaringan tubuh leukosit akan pindah menuju jaringan yang mengalami radang dengan cara menembus dinding kapiler (Kiswari, 2014).

2.4.6 Neutrophil

Neutrophil adalah bagian sel darah putih dari kelompok granulosit. *Neutrophil* berukuran sekitar 14 μm , granulanya berbentuk butiran halus tipis dengan sifat netral

sehingga terjadi pencampuran warna asam (eosin) dan warna basa (metilen biru), sedang pada granula menghasilkan warna ungu atau merah muda yang samar (Nugraha 2015). *Neutrophil* berfungsi sebagai garis pertahanan tubuh terhadap zat asing terutama terhadap bakteri. Bersifat fagosit dan dapat masuk ke dalam jaringan yang terinfeksi. Sirkulasi *neutrophil* dalam darah yaitu sekitar 10 jam dan dapat hidup selama 1-4 hari pada saat berada dalam jaringan ekstrasvaskuler. *Neutrophil* adalah jenis sel leukosit yang paling banyak yaitu sekitar 50-70% diantara sel leukosit yang lain. Ada dua macam netrofil yaitu *neutrophil* batang (stab) dan *neutrophil* segmen (polimorfonuklear) (Kiswari,2014).

Jumlah *neutrophil* yang rendah biasa disebut sebagai neutropenia. neutropenia merupakan salah satu efek samping dari kemoterapi yang disebabkan oleh supresi hematopoiesis yang menyebabkan neutropenia ditambah dengan faktor pencetus lain seperti infeksi. Neutropenia didefinisikan sebagai hitung absolut *neutrophil* kurang dari $500/\text{mm}^3$ dan dibagi menjadi 4 stadium menurut jumlah hitung absolutnya (Sharma & Lokeshwar, 2005). Penyebab neutropenia bermacam-macam. Neutropenia bisa terjadi akibat infeksi, obat-obatan, radioterapi, kemoterapi, hipersplenism, penggantian sumsum tulang, anemia, kegagalan sumsum tulang, anemia hipoplastik, defisiensi nutrisi, vitamin B12 dan defisiensi folat, serta neutropenia siklik yang merupakan salah satu penyakit periodik dengan infeksi berulang oleh karena adanya ritme di dalam tubuh akibat perubahan siklik dalam produksi dan pengeluaran *neutrophil* di sumsum tulang (Cameron, 2009) Neutropenia selektif dapat terjadi setelah terapi dengan sejumlah besar obat. Sebagian besar obat tersebut merusak prekursor sumsum tulang sehingga menghambat replikasi normal sel-sel di sumsum tulang dengan akibat neutropenia, trombositopenia, atau anemia (Sharma & Lokeshwar, 2005). Kemoterapi kanker dengan obat siklofosamid, nitrogen mustard, methotrexate, sitarabin, dan banyak lainnya dilaporkan sebagai penyebab neutropenia pada hampir 90% kasus.3 Dalam kemoterapi menggunakan obat-obat

tersebut, terjadi penurunan hematopoiesis oleh karena supresi myeloid (Heather & Gerald, 2008).

2.4.7 Indeks Massa Tubuh (IMT)

Indeks Massa Tubuh (IMT) atau Body Mass Indeks (BMI) adalah perbandingan (rasio) berat badan/tinggi badan yang sering digunakan untuk menilai berat badan orang dewasa, untuk mengetahui apakah berat badannya tergolong kurang, normal, lebih atau obese. IMT adalah berat badan dalam kilogram dibagi dengan tinggi badan kuadrat dalam meter. Rumus perhitungan Indeks Massa Tubuh dapat dilihat pada persamaan (2.45) sebagai berikut.

$$IMT = \frac{BB \text{ (Kg)}}{TB \text{ (m)} \times TB \text{ (m)}} \quad (2.45)$$

dimana,

IMT = Indeks Massa Tubuh

BB = Berat Badan

TB = Tinggi Badan

Menurut WHO (2007), IMT dikategorikan dengan hasil yang dapat dilihat pada Tabel 2.2 Sebagai berikut

Tabel 2.2 Klasifikas IMT

Indeks massa tubuh	Kategori IMT
<18,5	Berat badan kurang
18,5-24,9	Normal
25,0-29,9	Berat badan lebih
30,0-39,9	Obesitas
>40	Sangat obesitas

BAB III

METODOLOGI PENELITIAN

3.1 Sumber Data

Data yang digunakan dalam penelitian ini adalah data sekunder yang diperoleh dari penelitian yang dilakukan oleh Tamoki (2017). Data tersebut merupakan data pasien yang menjalani kemoterapi berbasis *5-FU* dari *specialist of gastrointestinal surgery* di Nippon Medical School (Universitas swasta di Tokyo, Jepang). Pada data tersebut, studi klinis dilakukan terhadap 28 pasien yang menjalani kemoterapi dengan *5-FU*. Setelah melakukan kemoterapi pada tubuh pasien dipasangkan *activity meter* untuk mengukur jumlah langkah kaki pasien dan jam tidur tidak nyenyak pasien yang digunakan sebagai variabel dalam penelitian ini.

3.2 Variabel Penelitian

Variabel yang digunakan dalam penelitian ini terdiri dari dua variabel, yaitu variabel respon (Y) dan variabel prediktor (X). Dalam melakukan analisis regresi multivariat digunakan 3 variabel respon yaitu variabel Y_1 , Y_2 dan Y_3 . Sedangkan pada analisis regresi logistik biner menggunakan 1 variabel respon yaitu Y_4 . Pada penelitian ini untuk analisis regresi multivariat dan regresi logistik akan digunakan variabel prediktor yang sama. Variabel respon yang digunakan pada penelitian ini akan ditampilkan pada Tabel 3.1 sebagai berikut.

Tabel 3.1 Variabel Respon Penelitian

	Variabel	Deskripsi	Jenis data
Y_1	<i>White Blood Cells</i> (WBC)	Perubahan Jumlah <i>White Blood Cells</i> (WBC) (sel/ μ l)	Rasio
Y_2	<i>Neutrophil</i>	Perubahan <i>Neutrophil</i> dalam WBC (sel/ μ l)	Rasio
Y_3	Rata-rata Jam tidur tidak nyenyak (Shalow Sllap)	Rata-rata jam tidur tidak nyenyak pasien selama 14 hari setelah melakukan kemoterapi (menit/hari)	Rasio

Tabel 3.1 Variabel Respon Penelitian (lanjutan)

Variabel	Deskripsi	Jenis data
Y ₄ Total WBC Count	0 = Jumlah WBC setelah kemoterapi menurun 1 = Jumlah WBC setelah kemoterapi meningkat	Nominal

Variabel prediktor yang digunakan pada penelitian ini ditunjukkan pada Tabel 3.2 sebagai berikut.

Tabel 3.2 Variabel Prediktor Penelitian

Variabel	Deskripsi	Jenis data
X ₁ Jenis Kelamin	0 = Laki-laki 1 = Perempuan	Nominal
X ₂ Umur	Umur pasien (tahun)	Rasio
X ₃ Indeks Massa Tubuh	Perbandingan berat badan/tinggi badan 1 = IMT <18,5 2 = IMT 18,5-24,9 3 = IMT 25,0-29,9 4 = IMT 30-39,9 5 = IMT >40	Nominal
X ₄ Rata-rata Jumlah Langkah Kaki	Rata-rata jumlah langkah kaki pasien selama 14 hari setelah melakukan kemoterapi (langkah/hari)	Rasio
X ₅ 5-FU 1	Dosis senyawa anti kanker 5-FU 1 yang digunakan untuk kemoterapi	Rasio
X ₆ 5-FU 2	Dosis senyawa ant kanker 5-FU 2 yang digunakan untuk kemoterapi	Rasio
X ₇ LV (<i>levamisole</i>)	Dosis senyawa <i>Levamisole</i>	Rasio
X ₈ L-OHP (<i>Oxaliplatin</i>)	Dosis senyawa L-OHP (<i>Oxaliplatin</i>)	Rasio
X ₉ IRI (<i>Irinotecan</i>)	Dosis senyawa IRI (<i>Irinotecan</i>)	Rasio
X ₁₀ B (<i>Bevacizumab</i>)	Dosis senyawa B (<i>Bevacizumab</i>)	Rasio

Tabel 3.2 Variabel Prediktor Penelitian (lanjutan)

Variabel	Deskripsi	Jenis data
X_{11} C (<i>Capecitabine</i>)	Dosis senyawa C (<i>Capecetabine</i>)	Rasio
X_{12} P (<i>Panitumumab</i>)	Dosis senyawa P (<i>Panitumumab</i>)	Rasio

3.3 Struktur Data

Struktur data yang akan dianalisis menggunakan regresi multivariat ditunjukkan pada Tabel 3.3 sebagai berikut.

Tabel 3.3 Struktur Data Regresi Multivariat

Observasi	Respon			Prediktor			
	Y_1	Y_2	Y_3	X_1	X_2	...	X_{12}
1	$Y_{1,1}$	$Y_{2,1}$	$Y_{3,1}$	$X_{1,1}$	$X_{2,1}$...	$X_{12,1}$
2	$Y_{1,2}$	$Y_{2,2}$	$Y_{3,2}$	$X_{1,2}$	$X_{2,2}$...	$X_{12,2}$
3	$Y_{1,3}$	$Y_{2,3}$	$Y_{3,3}$	$X_{1,3}$	$X_{2,3}$...	$X_{12,3}$
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
28	$Y_{1,28}$	$Y_{2,28}$	$Y_{3,28}$	$Y_{1,28}$	$Y_{2,28}$...	$Y_{12,28}$

Struktur data yang akan dianalisis menggunakan regresi logistik biner ditunjukkan pada Tabel 3.4 sebagai berikut.

Tabel 3.4 Struktur Data Regresi Logistik Biner

Observasi	Respon	Prediktor			
	Y_4	X_1	X_2	...	X_{12}
1	$Y_{4,1}$	$X_{1,1}$	$X_{2,1}$...	$X_{8,1}$
2	$Y_{4,2}$	$X_{1,2}$	$X_{2,2}$...	$X_{8,2}$
3	$Y_{4,3}$	$X_{1,3}$	$X_{2,3}$...	$X_{8,3}$
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
28	$Y_{4,28}$	$Y_{1,28}$	$Y_{2,28}$...	$Y_{12,28}$

3.4 Langkah Penelitian

Langkah-langkah analisis yang akan dilakukan berdasarkan dengan tujuan pada penelitian ini adalah sebagai berikut.

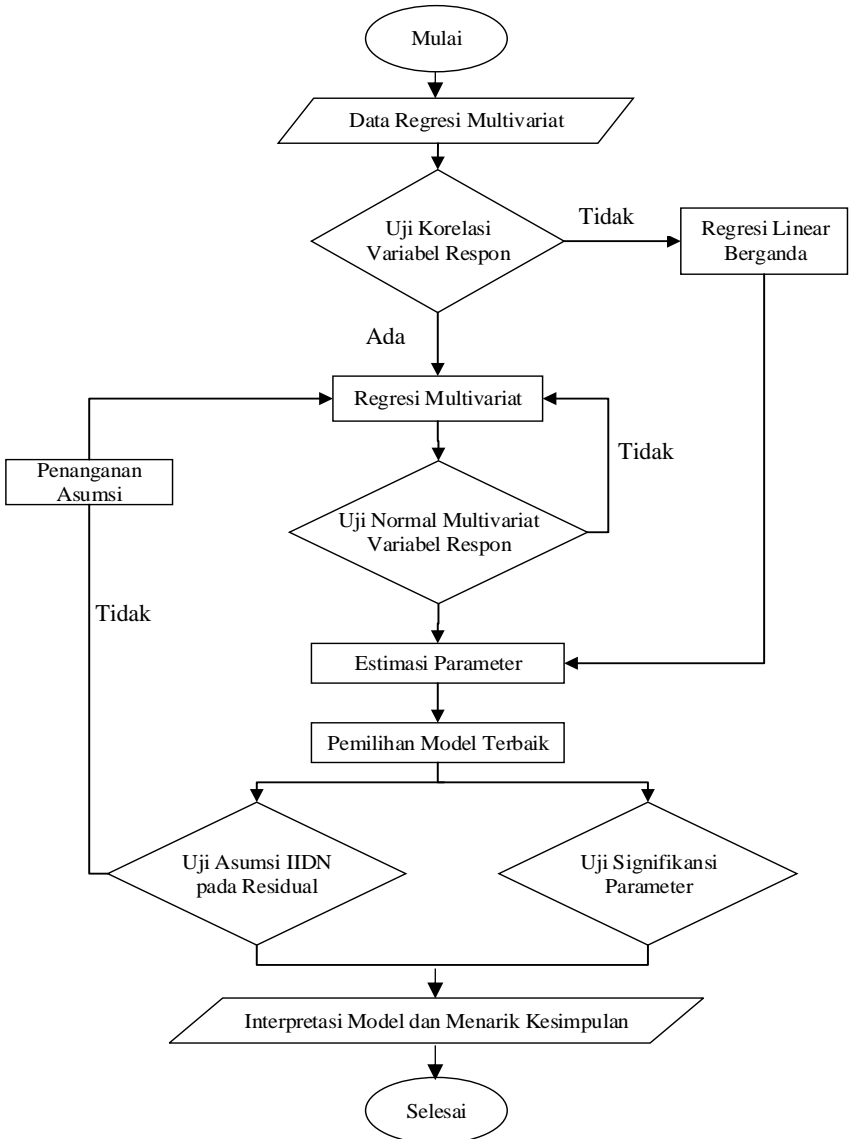
1. Eksplorasi data dan mendeskripsikan faktor-faktor yang diduga mempengaruhi keberhasilan pengobatan kemoterapi menggunakan senyawa 5-FU
2. Melakukan analisis regresi multivariat untuk mengetahui faktor-faktor yang mempengaruhi keberhasilan pada

kemoterapi. Pada pengujian multivariat, variabel response yang digunakan adalah perubahan jumlah WBC (Y_1), perubahan jumlah *Neutrophil* (Y_2) dan rata-rata jam tidur tidak nyenyak (Y_3), dengan langkah-langkah sebagai berikut.

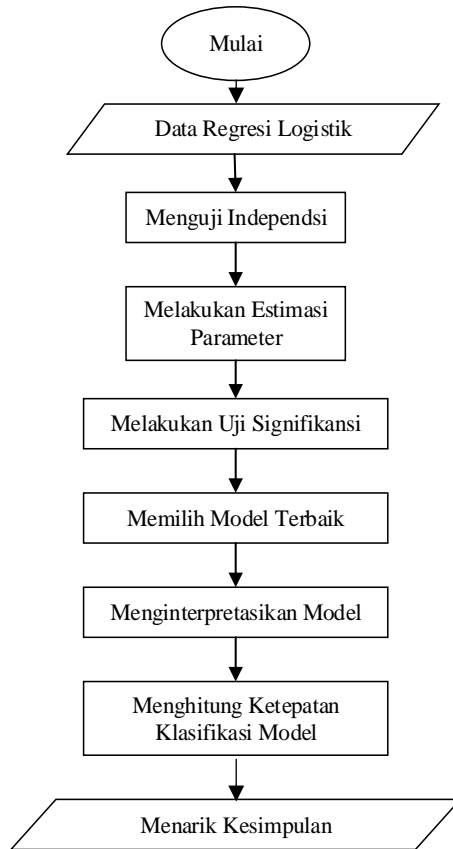
- a. Melakukan pengujian dependensi antar variabel respon dengan menggunakan *Bartlett's Test*. Jika terbukti variabel respon berkorelasi maka analisis dapat dilanjutkan pada tahap selanjutnya dengan metode regresi multivariat.
 - b. Melakukan pengujian berdistribusi normal multivariat pada semua variabel respon.
 - c. Melakukan pembentukan estimasi parameter model regresi multivariat
 - d. Melakukan pengujian signifikansi parameter model, meliputi pengujian signifikansi parameter model regresi multivariat secara serentak dan parsial.
 - e. Melakukan pengujian asumsi residual IIDN secara multivariat.
 - f. Menghitung nilai ukuran kebaikan model antar variabel respon dan variabel prediktor.
 - g. Mendapatkan model persamaan regresi multivariat untuk mengukur faktor-faktor yang mempengaruhi keberhasilan kemoterapi .
3. Melakukan analisis regresi secara univariat menggunakan regresi linear berganda untuk mengetahui faktor-faktor yang mempengaruhi masing-masing variabel respon yaitu, perubahan *White Blood Cells* (Y_1), perubahan *Neutrophil* (Y_2) dan rata-rata jam tidur tidak nyenyak (Y_3), dengan langkah-langkah sebagai berikut.
- a. Melakukan estimasi parameter regresi linear berganda dengan metode *Stepwise* pada masing-masing variabel respon
 - b. Melakukan pengujian signifikansi parameter secara serentak pada masing-masing variabel respon
 - c. Melakukan pengujian signifikansi parameter secara parsial pada masing-masing variabel respon

- d. Pengujian residual regresi identik, independen, berdistribusi normal (IIDN) dan uji multikolinearitas pada masing-masing variabel respon.
- 4. Melakukan analisis regresi logistik biner, dimana variabel respon yang digunakan berupa perubahan *total white blood cells count* pasien setelah melakukan kemoterapi. Nilai *total white blood cells count* dikategorikan menjadi dua yaitu perubahan jumlah WBC bertambah dan perubahan jumlah WBC berkurang. Langkah-langkah melakukan analisis regresi logistik biner adalah sebagai berikut.
 - a. Melakukan uji independensi antara variabel respon (Y) dan variabel prediktor (X) untuk mengetahui hubungan antara variabel respon (Y) dan variabel prediktor (X)
 - b. Melakukan estimasi Parameter model regresi logistik
 - c. Melakukan uji signifikansi parameter secara serentak dan parsial
 - d. Memilih model terbaik menggunakan metode *Bacward Stepwise (Wald)*. Kemudian menginterpretasikan model yang diperoleh berdasarkan *oods ratio* dan nilai peluang dari regresi logistik
 - e. Menghitung nilai akurasi dari hasil klasifikasi model untuk mengetahui seberapa besar observasi secara tepat diklasifikasikan
- 5. Menarik kesimpulan dan saran dari hasil analisis untuk masing-masing metode

Berdasarkan langkah-langkah penelitian yang akan dilakukan dan penjelasan sebelumnya, pada penelitian ini akan menggunakan tiga diagram alir. Diagram alir untuk analisis regresi multivariat dapat dilihat pada Gambar 3.1 dan diagram alir untuk analisis regresi logistik biner dapat dilihat pada gambar 3.2



Gambar 3.1 Diagram Alir Regresi Multivariat



Gambar 3.2 Diagram Alir Regresi Logistik Biner

(Halaman ini sengaja dikosongkan)

BAB IV

ANALISIS DAN PEMBAHASAN

Pada bab ini dibahas mengenai hasil analisis dan pembahasan berdasarkan pengolahan data yang telah dilakukan. Analisis yang dilakukan meliputi statistika deskriptif untuk menggambarkan karakteristik kasus kemoterapi kanker serta analisis regresi multivariat dan regresi logistik biner untuk mengetahui faktor-faktor yang mempengaruhi keberhasilan kemoterapi kanker.

4.1 Karakteristik Data Pasien Kemoterapi Kanker

Karakteristik data pasien kemoterapi kanker perlu dikaji untuk memperoleh informasi secara umum tentang variabel yang digunakan dalam penelitian, diharapkan informasi tersebut dapat mempermudah untuk mengetahui faktor-faktor yang mempengaruhi keberhasilan kemoterapi kanker.

Tabel 4.1 merupakan deskripsi data untuk variabel respon yang digunakan pada analisis regresi multivariat. Dimana variabel respon yang digunakan pada analisis regresi multivariat digunakan tiga variabel respon yaitu variabel perubahan *White Blood Cells* (Y_1), perubahan *Neutrophil* (Y_2) dan rata-rata jam tidur tidak nyenyak pasien (Y_3).

Tabel 4.1 Statistika Deskriptif Variabel Respon

Variabel		Rata-rata	Standar Deviasi	Minimum	Maksimum
Y_1	Meningkat	869,00	528,00	100,00	2000,00
	Menurun	-933,00	600,00	-2200,00	-100,00
Y_2	Meningkat	730,00	401,00	148,00	1210,00
	Menurun	-698,20	420,60	-1479,30	-39,40
Y_3		224,10	74,20	104,50	346,90

Tabel 4.1 menjelaskan bahwa rata-rata perubahan *White Blood Cells* pada pasien penderita kanker setelah melakukan kemoterapi cenderung mengalami penurunan sebesar 933 sel/ μ l dan rata-rata yang cenderung mengalami peningkatan sebesar 869 sel/ μ l, dengan standar deviasi masing masingnya adalah 600 sel/ μ l dan 528 sel/ μ l. Hal yang sama juga terjadi pada perubahan *neutrophil* setelah melakukan kemoterapi yang rata-rata mengalami penurunan sebesar -698,2 sel/ μ l dan peningkatan sebesar 730 sel/ μ l

dengan standar deviasi masing-masingnya 420,6 sel/ μ l dan 401 sel/ μ l. Standar deviasi yang besar menjelaskan bahwa data pada variabel perubahan WBC dan *Neutrophil* bersifat heterogen. Pada data jam tidur tidak nyenyak pasien, rata-rata jam tidur tidak nyenyak setelah melakukan kemoterapi adalah sebesar 224,1 menit setiap harinya dengan standar deviasi 74,2.

Selanjutnya pada Tabel 4.2 akan disajikan karakteristik data untuk variabel preditor yang memiliki skala rasio. Variabel prediktor yang digunakan meliputi umur (X_2), rata-rata jumlah langkah kaki pasien per hari (X_4), dosis obat kemoterapi 5-FU 1 (X_5), dosis obat kemoterapi 5-FU 2 (X_6) dan dosis obat kemoterapi *levamisole* (X_7)

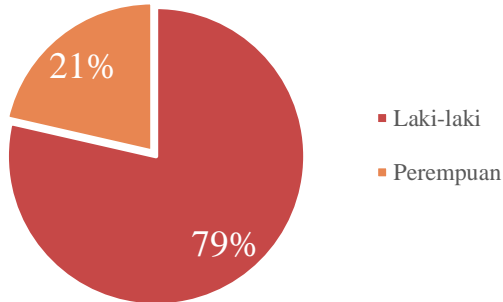
Tabel 4.2 Statistika Deskriptif Variabel Prediktor

Variabel	Rata-rata	Standar Deviasi	Minimum	Maximum
X_2	64,96	11,06	39,00	84,00
X_4	5333,00	2243,00	1753,00	10726,00
X_5	641,60	80,90	450,00	750,00
X_6	3823,20	480,80	2800,00	4650,00
X_7	327,50	31,02	250,00	380,00

Tabel 4.2 menjelaskan bahwa untuk rata-rata pasien yang melakukan kemoterapi mempunyai umur 64,96 tahun, dengan standar deviasi 11,06. umur minimum dari pasien kemoterapi adalah 39 tahun da umur *maximum* 84 tahun. Pada data pasien 14 hari setelah melakukan kemoterapi, rata-rata jumlah langkah kaki pasien adalah 5333 langkah perhari. Rata-rata dosis yang digunakan untuk senyawa kemoterapi kanker 5-FU 1 , 5-FU 2 dan LV masing masingnya adalah 641,6, 3823,2 dan 327,5. Dimana dosis *maximum* yang digunakan untuk senyawa kemoterapi kanker 5-FU 1 , 5-FU 2 dan LV masing masingnya adalah yaitu sebesar 750, 4650 dan 380. Serta dosis minimum untuk masing-masingnya adalah 450, 2800, dan 250.

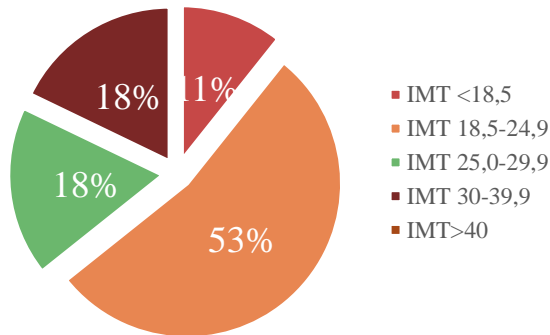
Selanjutnya akan divisualisasikan proporsi untuk variabel prediktor. Proporsi masing-masing variabel prediktor akan ditampilkan dalam bentuk *pie chart* yang digunakan pada penelitian ini diantaranya, variabel jenis kelamin pasien kemoterapi (X_1), indeks massa tubuh pasien (X_3), serta

menggunakan atau tidaknya senyawa kanker *Oxaliplatin* (X_8), *Irinotecan* (X_9), *Bevacizumab* (X_{10}), *Capecetabine* (X_{11}) dan *Panitumumab* (X_{12})



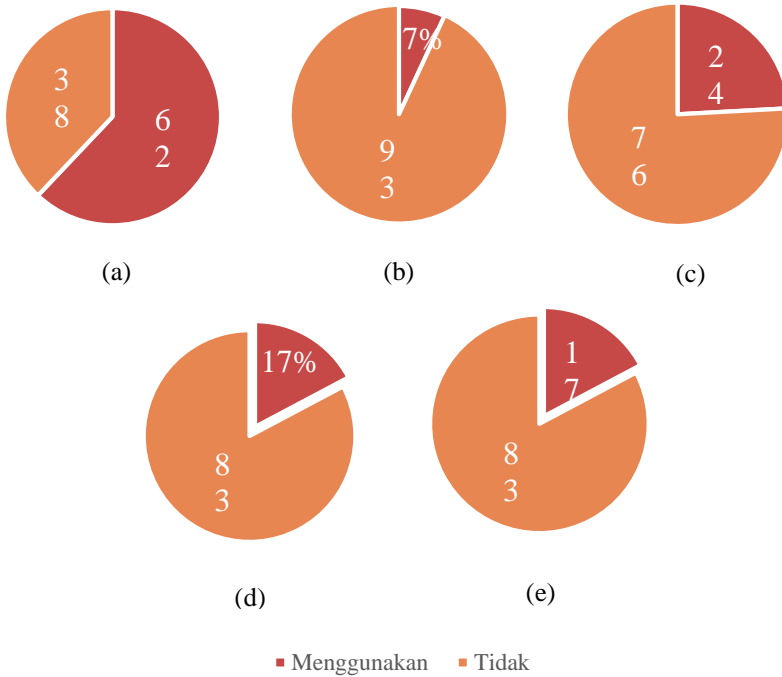
Gambar 4.1 Proporsi Pasien Kemoterapi Kanker Berdasarkan Jenis Kelamin

Gambar 4.1 menunjukkan bahwa dari data pasien kemoterapi kanker, jumlah pasien berjenis kelamin laki-laki lebih banyak daripada pasien perempuan. Pasien berjenis kelamin laki laki adalah sebanyak 79% dari dari total pasien dan jumlah pasien kemoterapi kanker yang berjenis kelamin perempuan adalah sebanyak 21%.



Gambar 4.2 Proporsi Pasien Kemoterapi Kanker Berdasarkan Indeks Massa Tubuh

Gambar 4.2 menjelaskan bahwa 53% dari pasien yang melakukan kemoterapi memiliki nilai indeks massa tubuh 18,5–24,9. Sedangkan sisanya, 11 % dari total pasien memiliki IMT kurang dari 18,5. dan masing-masing 18 % memiliki IMT sebesar 25,0 - 29,9 dan 30 – 39,9.



Gambar 4.3 Proporsi Pasien Menggunakan Senyawa Lain Saat Kemoterapi

Keterangan :

(a) : Senyawa *Oxaliplatin*

(b) : Senyawa *Irinotecan*

(c) : Senyawa *Bevacizumab*

(d) : Senyawa *Capecitabine*

(e) : Senyawa *Panitumumab*

Gambar 4.3 dapat dilihat bahwa dari total pasien yang melakukan kemoterapi menggunakan senyawa *Oxaliplatin* (L-OHP) 62% diantaranya melakukan kemoterapi menggunakan senyawa *Oxaliplatin* (L-OHP), sedangkan 38% diantaranya tidak menggunakan senyawa *Oxaliplatin*. Pasien yang melakukan kemoterapi menggunakan senyawa *Irinotecan* (IRI) sebanyak 7%

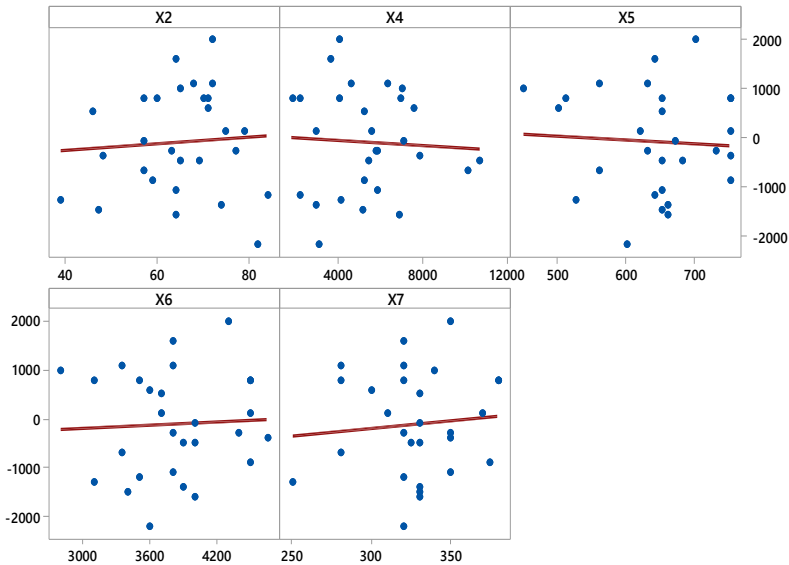
dari total semua pasien dan 93% diantaranya melakukan kemoterapi tidak menggunakan senyawa *Irinotecan* (IRI),

Jumlah pasien yang tidak menggunakan senyawa B (*Bevacizumab*) lebih banyak dibandingkan jumlah pasien yang melakukan kemoterapi menggunakan senyawa B (*Bevacizumab*). Dari seluruh total pasien, 24% diantaranya melakukan kemoterapi menggunakan senyawa B (*Bevacizumab*) dan 76% lainnya tidak menggunakan senyawa B (*Bevacizumab*) saat melakukan kemoterapi kanker. 17% diantara total pasien menggunakan senyawa *Capecitabine* pada saat melakukan kemoterapi dan 83% dari total pasien yang melakukan kemoterapi kanker tidak menggunakan senyawa *Capecitabine* pada proses pengobatan kanker melalui kemoterapi.

Jumlah pasien yang menggunakan senyawa *Panitumumab* saat melakukan kemoterapi lebih sedikit dibandingkan jumlah pasien yang tidak menggunakan senyawa tersebut. Dimana, dari total keseluruhan pasien, 83% diantaranya melakukan kemoterapi kanker tidak menggunakan senyawa *Panitumumab* pada proses pengobatan kanker melalui kemoterapi. Sedangkan 17% diantara total pasien menggunakan senyawa *Panitumumab*.

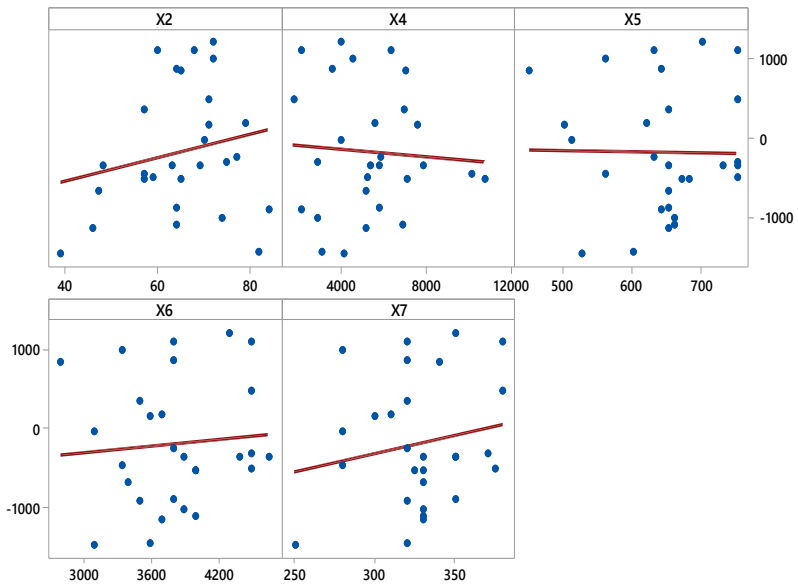
4.2 Pemodelan Menggunakan Regresi Multivariat

Pada regresi multivariat, sebelum melakukan pemodelan terlebih dahulu dilihat sebaran data serta pola hubungan linear antar variabel. serta akan dilihat besar korelasi antar variabel. Pada *scatterplot* akan dilihat pola hubungan antar variabel prediktor yang mempunyai skala rasio dengan masing-masing variabel prediktor yaitu variabel perubahan nilai total WBC, perubahan nilai total *neutrophil* dan rata-rata jam tidur tidak nyenyak pasien setelah melakukan kemoterapi. Berikut akan disajikan *scatterplot* untuk masing masing variabel yang akan disajikan pada Gambar 4.4, Gambar 4.5 dan Gambar 4.6.



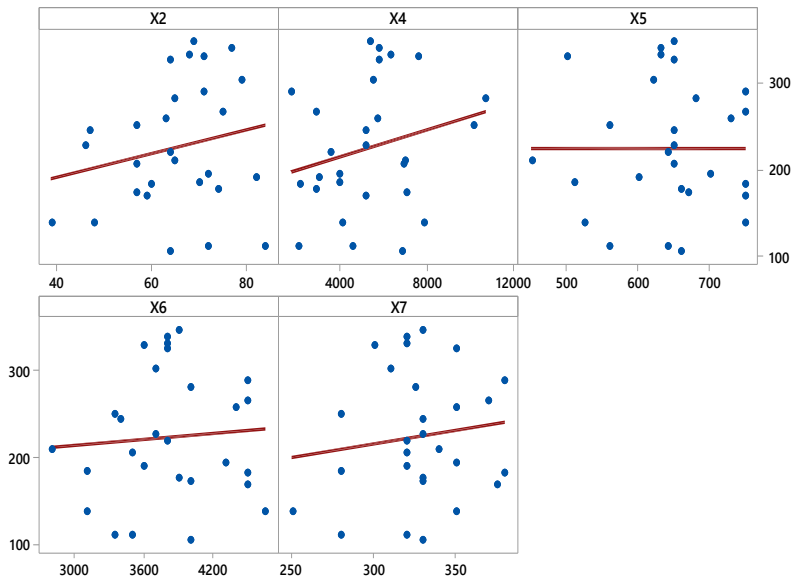
Gambar 4.4 Hubungan Variabel prediktor terhadap Variabel Y_1

Gambar 4.4 menjelaskan bahwa pada perubahan WBC (*White Blood Cells*) setelah kemoterapi terdapat pola hubungan berbanding lurus antara variabel umur pasien (X_2), dosis senyawa *5-Fluorouracil* 2 (X_6), dosis senyawa *levamisole* (X_7) terhadap perubahan WBC. Dimana, apabila masing-masing variabel umur pasien, dosis senyawa *5-FU* 2 atau dosis senyawa LV meningkat maka perubahan WBC atau sel darah putih juga meningkat. Sedangkan pada variabel rata-rata jumlah langkah kaki (X_4) dan senyawa *5-Fluorouracil* 1 (X_5) memiliki pola hubungan yang berbanding terbalik dengan perubahan WBC. Dimana, apabila jumlah langkah kaki atau *5-FU* 1 meningkat maka perubahan sel darah putih menurun. Selain itu, dapat dilihat pula bahwa data cenderung menyebar dan tidak membentuk pola tertentu seperti lingkaran atau kurva, sehingga dapat dikatakan bahwa data linear.



Gambar 4.5 Hubungan Variabel Prediktor terhadap Variabel Y_2

Pola hubungan antara masing-masing variabel prediktor terhadap perubahan *neutrophil* setelah melakukan kemoterapi dapat dilihat pada Gambar 4.9. Perubahan *Neutrophil* setelah melakukan kemoterapi memiliki pola hubungan berbanding lurus antara variabel umur pasien (X_2), 5-Fluorouracil 2 (X_6), dan dosis senyawa *levamisole* (X_7). Dimana, apabila umur pasien, dosis senyawa 5-FU 2 atau dosis senyawa LV meningkat maka perubahan *neutrophil* juga meningkat. Sedangkan pada variabel rata-rata jumlah langkah kaki (X_4) dan dosis 5-Fluorouracil 1 (X_5) memiliki pola hubungan yang berbanding terbalik dengan perubahan *neutrophil*. Dimana, apabila jumlah langkah kaki dan dosis senyawa 5-Fluorouracil 1 bertambah maka perubahan *neutrophil* akan menurun. Terlihat juga bahwa data cenderung menyebar dan tidak membentuk pola tertentu sehingga dapat disimpulkan pula bahwa data linear.



Gambar 4.6 Hubungan Variabel prediktor terhadap Variabel Y_3

Gambar 4.6 melihat pola hubungan antara lima variabel prediktor dengan jam tidur tidak nyenyak pasien. Secara visual dapat dilihat bahwa umur pasien (X_2), rata-rata jumlah langkah kaki (X_4), dosis senyawa 5- *Fluorouracil* 2 (X_6) dan dosis senyawa *levamisole* (X_7) yang digunakan dalam kemoterapi, memiliki pola hubungan yang berbanding lurus dengan jam tidur tidak nyenyak pada pasien. Dimana apabila bahwa variabel tersebut mengalami peningkatan, maka jam tidur tidak nyenyak pada pasien juga meningkat. Selain itu, penyebaran data yang tidak merata pada Gambar 4.10 membuktikan bahwa data tidak linear.

Selain menggunakan *scatterplot*, untuk melihat pola hubungan antar variabel dapat dilihat menggunakan nilai korelasi antar variabel respon dan variabel prediktor yang bersifat numerik. Korelasi antar variabel dapat dilihat pada Tabel 4.3.

Tabel 4.3 Korelasi Antar Variabel

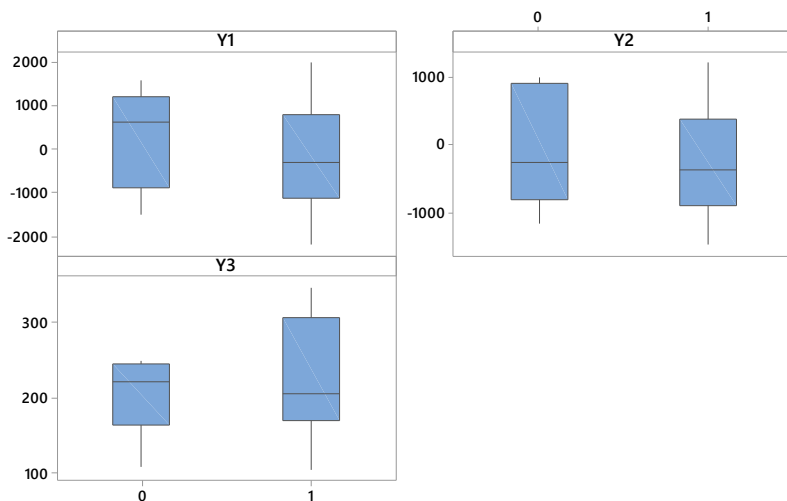
	Y ₁	Y ₂	Y ₃	X ₂	X ₄	X ₅
Y ₂	0,884					
Y ₃	0,154	0,163				
X ₂	0,07	0,205	0,206			
X ₄	-0,06	-0,066	0,232	-0,327		
X ₅	-0,061	-0,012	0,001	-0,046	-0,224	
X ₆	0,045	0,088	0,077	0,041	-0,153	0,922
X ₇	0,088	0,181	0,134	0,099	-0,254	0,776
X ₈	0,245	0,351	0,184	0,132	0,027	0,352
X ₉	-0,33	-0,286	-0,035	-0,025	0,127	0,047
X ₁₀	-0,185	-0,243	0,12	-0,27	-0,047	0,184
X ₁₁	-0,019	-0,139	-0,245	-0,372	0,098	0,221
X ₁₂	0,104	0,173	-0,312	0,324	-0,478	0,086

Tabel 4.3 Korelasi Antar Variabel (lanjutan)

	X ₆	X ₇	X ₈	X ₉	X ₁₀	X ₁₁
X ₇	0,773					
X ₈	0,444	0,31				
X ₉	0,045	0,114	-0,364			
X ₁₀	0,158	0,324	-0,322	0,415		
X ₁₁	0,196	0,024	0,08	-0,128	-0,263	
X ₁₂	0,086	0,142	-0,223	-0,126	-0,258	-0,209

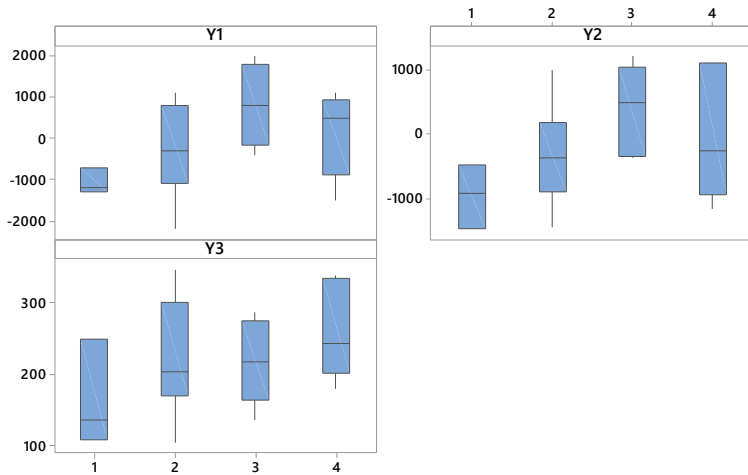
Nilai koefisien korelasi antar variabel dapat dilihat pada Tabel 4.3 yaitu dapat dilihat bahwa semakin tinggi nilai yang ditunjukkan, maka semakin besar korelasi atau hubungan antar variabel. Dimana, apabila nilai korelasi bernilai negatif menunjukkan hubungan berbanding terbalik antar variabel, sedangkan nilai positif menunjukkan hubungan berbanding lurus antar variabel.

Pada variabel prediktor yang bersifat kategorik, untuk melihat perbedaan pengaruh antar kategorik variabel dengan variabel respon dapat dilihat menggunakan *boxplot*. Perbedaan pengaruh antara kategorik variabel prediktor akan terhadap masing-masing variabel respon yang digunakan dapat dilihat dari *boxplot* berdasarkan masing-masing kategorik variabel prediktor. Berikut merupakan *boxplot* tiap variabel prediktor yang bersifat kategorik dengan masing masing variabel respon pada Gambar 4.7 dan Gambar 4.8.



Gambar 4.7 Variabel Respon Berdasarkan Jenis Kelamin

Pasien berjenis kelamin laki-laki (1) dan perempuan (2) tidak memberikan efek yang berbeda terhadap perubahan *White Blood Cells*, *Neutrophil* serta rata-rata jam tidur tidak nyenyak pasien setelah melakukan kemoterapi. Hal tersebut terlihat pada Gambar 4.7, dari bentuk *boxplot* antara kategorik 0 (pasien berjenis kelamin perempuan) dan 1 (pasien berjenis kelamin laki-laki) memiliki perpotongan atau perserbaran data yang sama antara kategorik 0 dan 1 pada ketiga variabel respon.



Gambar 4.8 Variabel Respon Berdasarkan Indeks Massa Tubuh

Gambar 4.8 menunjukkan bahwa pada Gambar (a) dan Gambar (b) terdapat persebaran data yang berbeda antar kategori 1 dan kategori 3 atau tidak terjadi perpotongan antara *Boxplot* kategori 1 dan kategori 3. Sehingga dapat disimpulkan bahwa indeks massa tubuh pasien memberikan efek yang berbeda terhadap perubahan total WBC dan perubahan *neutrophil*. Sedangkan pada Gambar (c) pada *boxplot* ke empat kategori terjadi perpotongan atau persebaran data cenderung terlihat sama, sehingga dapat disimpulkan bahwa indeks massa tubuh pasien memberikan efek yang berbeda pada tiap kategoriknya terhadap rata-rata jam tidur tidak nyenyak pasien.

4.2.1 Pengujian Dependensi Antar Variabel Respon

Pengujian yang sesuai untuk mengetahui kebebasan antar variabel respon dapat dilakukan dengan menggunakan uji Barlett Sphericity. Dengan hipotesis sebagai berikut :

$H_0 : \rho = I$ (matriks korelasi sama dengan matriks identitas atau data indepeneden)

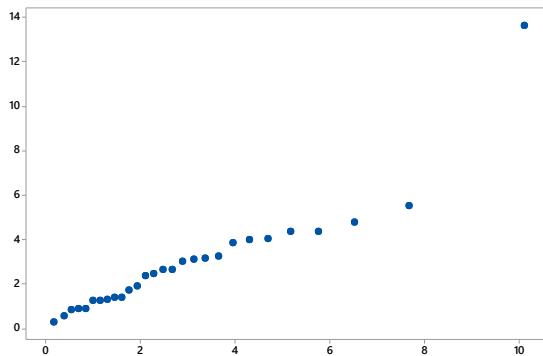
$H_1 : \rho \neq I$ (matriks korelasi tidak sama dengan matriks identitas atau data dependen)

Pengujian dilakukan dengan membandingkan nilai χ^2 dengan $\chi^2_{\alpha; \frac{1}{2}q(1-q)}$ yang merupakan daerah kritis. Keputusan H_0 ditolak jika $\chi^2_{hitung} \geq \chi^2_{\alpha; \frac{1}{2}q(1-q)}$ yang berarti antar variabel respon bersifat *dependent*. Nilai q merupakan jumlah variabel respon dan $\ln|R|$ merupakan determinan matrik korelasi dari masing-masing variabel respon.

Dari hasil perhitungan berdasarkan Persamaan (2.3) diperoleh nilai χ^2_{hitung} sebesar 38,892 dengan nilai p -value sebesar 0,00. Dengan nilai tabel kritis yaitu $\chi^2_{0,05;3}$ sebesar 7,815, maka dapat diambil keputusan bahwa H_0 ditolak karena nilai $\chi^2_{hitung} > \chi^2_{\alpha; \frac{1}{2}q(1-q)}$ atau nilai p -value lebih kecil dari nilai signifikansi (α) 0,05. Sehingga dapat disimpulkan bahwa antara variabel respon saling berkorelasi atau *dependent* dengan taraf signifikansi 5%.

4.2.2 Pengujian Normal Multivariat

Normal multivariat pada variabel respon dilakukan dengan koefisien korelasi plot *chi-square*. Berikut merupakan qq plot yang didapatkan dari variabel respon yang digunakan.



Gambar 4.9 QQ-plot Variabel Respon

Gambar 4.9 menunjukkan sebaran data cenderung menyerupai garis lurus atau linear jadi dapat disimpulkan bahwa data berdistribusi normal multivariat. Kesimpulan diperkuat dengan hasil perhitungan yang didapatkan proporsi

$$d_i^2 < \chi_p^2 \left(\left(n - i + \frac{1}{2} \right) / n \right) \text{ sebesar } 46,45\% \text{ atau nilai proporsi}$$

mendekati 50%. Dari nilai proporsi juga dapat disimpulkan bahwa variabel respon mengikuti distribusi normal multivariat.

4.2.3 Estimasi Parameter

Setelah dilakukan uji kebebasan antar variabel respon, deteksi multikolinearitas dan asumsi variabel respon berdistribusi normal multivariat, didapatkan hasil bahwa variabel respon telah memenuhi asumsi *dependentt* dan berdistribusi normal multivariat.. Kemudian dilakukan estimasi parameter untuk mendapatkan model persamaan regresi multivariat. Dalam analisis regresi multivariat ditentukan estimasi parameter regresi dengan menggunakan rumus persamaan (2.12).

Berikut ini merupakan model persamaan regresi multivariat yang terbentuk.

$$\begin{aligned} \hat{Y}_1 = & 2026,273 + 345,73 X_{1_0} - 10,27 X_2 - 1424,95 X_{3_1} - 408,95 \\ & X_{3_2} + 987,962 X_{3_3} + 0,03 X_4 - 8,61 X_5 + 0,53 X_6 + 6,78 X_7 \\ & + 0,32 X_8 - 2,81 X_9 - 1,72 X_{10} - 0,42 X_{11} + 0,416 X_{12} \end{aligned}$$

$$\begin{aligned} \hat{Y}_2 = & 1858,73 - 157,06 X_{1_0} - 4,79 X_2 - 1105,42 X_{3_1} - 389,42 \\ & X_{3_2} + 722,10 X_{3_3} + 0,49 X_4 - 3,20 X_5 - 0,36 X_6 + 4,48 X_7 + \\ & 3,47 X_8 - 0,85 X_9 - 1,00 X_{10} - 0,59 X_{11} + 1,14 X_{12} \end{aligned}$$

$$\begin{aligned} \hat{Y}_3 = & 164,79 - 49,09 X_{1_0} + 2,30 X_2 - 96,16 X_{3_1} - 82,51 X_{3_2} - \\ & 31,86 X_{3_3} + 0,01 X_4 - 0,14 X_5 + 0,01 X_6 + 0,04 X_7 - 0,04 X_8 \\ & - 0,15 X_9 + 0,02 X_{10} - 0,06 X_{11} - 0,22 X_{12} \end{aligned}$$

4.2.4 Pengujian Signifikansi Serentak Secara Multivariat

Pengujian signifikan serentak dilakukan secara multivariat dan univariat. Secara multivariat dengan menggunakan uji Wilk's Lambda sedangkan untuk pengujian secara univariat dengan menggunakan uji F. Berikut ini hasil pengujian signifikan serentak secara multivariat dan univariat.

Pengujian secara serentak dilakukan bertujuan untuk mengetahui apakah secara keseluruhan parameter signifikan dalam model. Untuk pengujian signifikansi model dengan menggunakan uji Wilk's Lambda. Berikut adalah hipotesis yang digunakan dan hasil perhitungan yang didapatkan.

$$H_0 : \beta_1 = \beta_2 = \dots = \beta_{11} = \beta_{12} = 0$$

$$H_1 : \text{paling sedikit ada satu } \beta_{jk} \neq 0$$

$$\text{Statistik Uji: } \Lambda = \frac{|E|}{|E + H|} = 0,974$$

$$p\text{-value} = 0,96$$

Pada statistik uji didapatkan bahwa nilai Wilk's Lambda adalah sebesar 0,974 dengan *p-value* 0,96. dimana pada nilai tersebut dapat diambil keputusan gagal tolak H_0 dan dapat disimpulkan bahwa pada pengujian secara serentak multivariat tidak terdapat variabel yang signifikan terhadap model persamaan karena pada hasil perhitungan nilai *p-value* > taraf signifikansi (α) yaitu sebesar 0,05.

4.2.5 Pengujian Signifikansi Parsial Secara Multivariat

Setelah dilakukukan pengujian signifikansi secara serentak kemudian dilakukan pengujian signifikansi secara parsial yang bertujuan untuk melihat pengaruh signifikan masing-masing variabel prediktor terhadap variabel respon. Pengujian signifikan parsial dilakukan secara multivariat dengan menggunakan uji Wilk's Lambda dengan hipotesis sebagai berikut:

$$H_0 : \beta_j = 0, j = 1, 2, \dots, 12$$

$$H_1 : \beta_j \neq 0$$

$$\text{Statistik Uji : } \Lambda = \frac{|E|}{|E + H|}$$

Dengan menggunakan statistik uji nilai *Wilk's Lambda* didapatkan hasil perhitungan *Wilk's Lambda* yang dapat dilihat pada Tabel 4.4

Tabel 4.4 Pengujian Signifikansi Parsial Secara Multivariat

	Variabel	Λ_{hitung}	$p-value$	Keputusan
X _{1_0}	Jenis Kelamin (perempuan)	0,862	0,637	Gagal Tolak H ₀
X ₂	Umur	0,892	0,725	Gagal Tolak H ₀
X _{3_1}	Indeks Massa Tubuh (1)	0,759	0,367	Gagal Tolak H ₀
X _{3_2}	Indeks Massa Tubuh (2)	0,765	0,380	Gagal Tolak H ₀
X _{3_3}	Indeks Massa Tubuh (3)	0,906	0,768	Gagal Tolak H ₀
X ₄	Rata-rata Jumlah Langkah Kaki	0,884	0,702	Gagal Tolak H ₀
X ₅	5-FU 1	0,921	0,815	Gagal Tolak H ₀
X ₆	5-FU 2	0,932	0,849	Gagal Tolak H ₀
X ₇	LV	0,989	0,988	Gagal Tolak H ₀
X ₈	L-OHP (<i>Oxaliplatin</i>)	0,891	0,724	Gagal Tolak H ₀
X ₉	IRI (<i>Irinotecan</i>)	0,930	0,840	Gagal Tolak H ₀
X ₁₀	B (<i>Bevacizumab</i>)	0,976	0,963	Gagal Tolak H ₀
X ₁₁	C (<i>Capecitabine</i>)	0,879	0,686	Gagal Tolak H ₀
X ₁₂	P (<i>Panitumumab</i>)	0,881	0,691	Gagal Tolak H ₀

Daerah kritis : Tolak H₀ jika $\Lambda_{hitung} \leq \Lambda_{\alpha, q, p, n-p-1}$

Hasil perhitungan yang disajikan pada Tabel 4.4 dapat diambil keputusan gagal tolak H₀ pada taraf signifikansi 5%,. Sehingga dapat disimpulkan bahwa tidak terdapat variabel prediktor yang pada model yang berpengaruh signifikan secara multivariat terhadap perubahan *White Blood Cells*, perubahan *Neutrophil* dan rata-rata jam tidur tidak nyenyak pasien pada taraf signifikansi 5%.

Hasil pengujian signifikansi pada analisis regresi multivariat didapatkan kesimpulan bahwa tidak terdapat variabel prediktor yang berpengaruh signifikan terhadap variabel respon. Pola hubungan antara variabel prediktor terhadap variabel respon susah untuk diprediksi dikarenakan *range* antara nilai minimum dan maksimum dari variabel respon yang teralalu besar. Selanjutnya, untuk melihat pengaruh variabel prediktor terhadap masing-masing variabel respon tersebut secara univariat dapat dianalisis menggunakan metode Regresi Linear berganda.

4.3 Pemodelan Menggunakan Regresi Linear Berganda Menggunakan Metode *Stepwise*

Pemodelan menggunakan regresi linear berganda dilakukan menggunakan metode *stepwise* untuk mengetahui variabel prediktor yang berpengaruh secara univariat terhadap masing-masing variabel respon. Metode *stepwise* merupakan salah satu metode untuk mendapatkan model terbaik dari sebuah analisis regresi. Secara definisi metode *stepwise* adalah gabungan antara metode *forward* dan *backward*. Dimana, variabel yang pertama kali masuk adalah variabel yang korelasinya tertinggi dan signifikan dengan variabel *dependentt*, variabel yang masuk kedua adalah variabel yang korelasi parsialnya tertinggi dan masih signifikan, setelah variabel tertentu masuk ke dalam model maka variabel lain yang ada di dalam model dievaluasi, jika ada variabel yang tidak signifikan maka variabel tersebut dikeluarkan.

4.3.1 Model Regresi Linear Berganda pada Perubahan WBC

Pemodelan pada variabel perubahan *White Blood Cells* setelah melakukan kemoterapi didapatkan tabel ANOVA yang ditunjukkan pada Tabel 4.5.

Tabel 4.5 *Analysis of Variance* (ANOVA) Variabel Y_1

<i>Source</i>	<i>DF</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>P</i>
<i>Regression</i>	5	14127419	2825484	3,68	0,014
<i>Error</i>	22	16882224	767374		
<i>Total</i>	27	31009643			

Tabel 4.5 merupakan hasil *Analysis of Variance* (ANOVA) yang memberikan hasil bahwa dengan taraf signifikansi (α) sebesar 0,05 didapatkan *p-value* kurang dari taraf signifikansi. Sehingga dapat disimpulkan bahwa minimal ada satu variabel yang berpengaruh signifikan terhadap perubahan *White Blood Cells* setelah melakukan kemoterapi.

Setelah mengetahui bahwa terdapat minimal satu variabel yang berpengaruh signifikan terhadap perubahan *White Blood Cells* setelah melakukan kemoterapi, kemudian dilakukan pengujian parsial untuk mengetahui variabel mana saja yang berpengaruh signifikan terhadap perubahan *White Blood Cells* setelah melakukan kemoterapi.

Tabel 4.6 Estimasi dan Pengujian Parameter Model Regresi Linear Y_1

Prediktor	Coef	SE	T	P	VIF
<i>Constant</i>	2079	1515	1,37	0,184	
X_5	-5,47	2.48	-2,20	0,038	1,42
$X_{3,2}$	1185	571	2,07	0,050	2,96
$X_{3,3}$	2669	732	3,65	0,001	2,87
$X_{3,4}$	1663	675	2,46	0,022	2,44
X_9	-4,12	2.69	-1,53	0,139	1,09

Tabel 4.6 menjelaskan bahwa nilai VIF tidak ada yang melebihi 10. Hal ini menunjukkan bahwa pada model regresi menggunakan metode *stepwise* tidak terdapat multikolinearitas antara variabel prediktor yang dimasukkan kedalam model.

Kemudian pada pengujian signifikansi variabel dengan taraf signifikansi (α) sebesar 0,05 didapatkan bahwa terdapat dua variabel respon yang berpengaruh signifikan terhadap perubahan *White Blood Cells* setelah melakukan kemoterapi yaitu dosis senyawa 5-FU 1 dan Indeks Massa Tubuh pasien. Sedangkan variabel dosis senyawa IRI (*Irinotecan*) berpengaruh signifikan terhadap perubahan *White Blood Cells* setelah melakukan kemoterapi pada taraf signifikansi sebesar 0,2.

Persamaan regresi linear untuk mengetahui pengaruh variabel prediktor terhadap perubahan WBC setelah melakukan kemoterapi sebagai berikut.

$$\hat{Y}_1 = 2079 + 1185X_{3,2} + 2669 X_{3,3} + 1663 X_{3,4} - 5,47 X_5 - 4,12 X_9$$

Variabel Indeks Massa Tubuh Pasien Merupakan variabel yang bersifat kategorik, sehingga dari model tersebut dapat dibentuk model regresi linear untuk masing masing kategorik Indeks Massa Tubuh Pasien yang ditunjukkan pada Tabel 4.7.

Tabel 4.7 Persamaan regresi linear Y_1

IMT	Persamaan
1	$\hat{Y}_1 = 2079 - 5,47 X_5 - 4,12 X_9$
2	$\hat{Y}_1 = 3264 - 5,47 X_5 - 4,12 X_9$
3	$\hat{Y}_1 = 4748 - 5,47 X_5 - 4,12 X_9$
4	$\hat{Y}_1 = 3741 - 5,47 X_5 - 4,12 X_9$

Keterangan :

X_5 = dosis senyawa 5 FU-1

X_9 = dosis senyawa *Irinotecan*

Tabel 4.7 Menjelaskan bahwa Jika pasien memiliki Indeks Massa Tubuh $<18,5$ ($IMT=1$) maka perubahan *White Blood Cells* pada pasien dapat dimodelkan sebagai dengan persamaan $Y_1 = 2079 - 5,47 X_5 - 4,12 X_9$. Yang artinya bahwa Perubahan WBC dipengaruhi oleh dosis 5 FU 1 yang apabila dosis ditambahkan 1 satuan, maka akan menyebabkan penurunan pada WBC pasien sebesar 5,47 dan dosis *irinotecan* yang apabila dosis senyawa *irinotecan* ditambahkan 1 satuan, maka akan menyebabkan penurunan pada WBC pasien sebesar 4,12 sel/ μ l dengan nilai konstan 2079. Begitu juga pengaruh dengan nilai konsta yang berbeda seterusnya sesuai kategorik Indeks Massa Tubuh pasien.

Model regresi linear tersebut menghasilkan nilai koefisien determinasi (R^2) sebesar 45,56%. Nilai tersebut menjelaskan bahwa variabel senyawa 5 FU-1 dan *Irinotecan* mampu menjelaskan pada penurunan jumlah *White Blood Cells* pada pasien sebesar 45,56% Sedangkan, sisanya dijelaskan oleh variabel lain yang tidak masuk kedalam model.

4.3.2 Pengujian Asumsi Model Regresi Linear Perubahan WBC

Terdapat asumsi residual yang harus dipenuhi dalam pemodelan regresi linear yaitu residual bersifat identik, independen, dan berdistribusi normal. Pengujian Asumsi IIDN pada residual adalah sebagai berikut.

a Uji Identik

Uji Glejser adalah pengujian yang digunakan untuk mengetahui kesamaan varian dari residual untuk semua pengamatan pada model regresi linear atau apakah residual memenuhi asumsi identik atau tidak. Berdasarkan hasil Uji *Glejser* didapatkan nilai *p-value* sebesar 1,000. Kesimpulan yang didapatkan dari nilia tersebut adalah varian dari residual sudah memenuhi asumsi identik pada taraf signifikansi 5%, karena *p-value* lebih besar dari taraf signifikansi.

b Uji Independen

Pengujian independen dilakukan menggunakan uji *Durbin-Watson*. Nilai *Durbin-Watson* yang didapatkan adalah sebesar 2,46. Dengan nilai DU dan 4-DL masing masingnya adalah 1,65 dan 2,82, maka dapat disimpulkan bahwa residual memenuhi asumsi independen karena nilai *Durbin-Watson* (d) berada pada interval $d_U < d < 4 - d_L$

c Uji Distribusi Normal

Pengujian distribusi normal digunakan untuk melihat apakah residual bersifat normal atau tidak. Pengujian ini menggunakan uji *Kolmogorov-Smirnov*. Pada pengujian *Kolmogorov-Smirnov* didapatkan nilai p -value sebesar 0,108. Dengan taraf signifikansi (α) sebesar 0,05 dapat disimpulkan bahwa residual berdistribusi normal karena nilai p -value lebih dari taraf signifikansi.

4.3.3 Model Regresi Linear Berganda pada Perubahan *Neutrophil*

Pada pemodelan variabel perubahan *neutrophil* setelah melakukan kemoterapi didapatkan tabel ANOVA yang ditunjukkan pada Tabel 4.8.

Tabel 4.8 *Analysis of Variance* (ANOVA) variabel Y_2

Source	DF	SS	MS	F	P
<i>Regression</i>	1	2165629,00	2165629,00	3,65	0,067
<i>Error</i>	26	15408115,00	592620,00		
<i>Total</i>	27	17573744,00			

Tabel 4.8 merupakan hasil *Analysis of Variance* (ANOVA) yang memberikan hasil bahwa minimal ada satu variabel yang berpengaruh signifikan terhadap perubahan *Neutrophil* setelah melakukan kemoterapi pada taraf signifikansi (α) sebesar 0,1. hal tersebut dapat dilihat dari nilai p -value sebesar 0,067 dimana nilai tersebut kurang dari taraf signifikansi (α) 10%.

Setelah mengetahui bahwa terdapat minimal satu variabel yang berpengaruh signifikan terhadap perubahan *Neutrophil* setelah melakukan kemoterapi, kemudian dilakukan pengujian parsial untuk mengetahui variabel mana saja yang berpengaruh

signifikan terhadap perubahan *Neutrophil* setelah melakukan kemoterapi.

Tabel 4.9 Estimasi dan Pengujian Parameter Model Regresi Linear Y_2

Prediktor	Coef	SE	T	P	VIF
Constant	-553	240	-2,30	0,029	
X_8	4,13	2,16	1,91	0,067	1,00

Tabel 4.9 menunjukkan bahwa nilai VIF tidak lebih dari 10. Hal ini menunjukkan bahwa pada model regresi menggunakan metode *stepwise* tidak terdapat multikolinearitas antara variabel prediktor yang dimasukkan kedalam model.

Kemudian pada pengujian signifikansi variabel dengan taraf signifikansi (α) sebesar 0,1 didapatkan bahwa dosis senyawa *Oxalipatin* berpengaruh signifikan terhadap perubahan *Neutrophil* setelah melakukan kemoterapi. Kemudian akan dibentuk model persamaan regresi linear berganda yang terbentuk untuk mengetahui variabel prediktor yang berpengaruh terhadap perubahan *Neutrophil* setelah melakukan kemoterapi menggunakan metode *stepwise*.

$$\hat{Y}_2 = -553 + 4,13 X_9$$

Persamaan diatas menjelaskan bahwa setiap peningkatan satu satuan dosis senyawa Irinotecan dan variabel lainnya dianggap konstan, maka akan mengakibatkan peningkatan jumlah *Neutrophil* setelah melakukan kemoterapi sebesar 0,185 sel/ μ l.

Model regresi linear diatas menghasilkan nilai koefisien determinasi (R^2) sebesar 12,32% dan MSE sebesar 155241227. Artinya bahwa variabel dosis senyawa *irinotecan* dapat menjelaskan peningkatan *Neutrophil* setelah melakukan kemoterapi sebesar 12,32%, sedangkan sisanya dijelaskan oleh variabel lain yang tidak dimasukkan kedalam model

4.3.4 Pengujian Asumsi Model Regresi Linear Perubahan *Neutrophil*

Terdapat asumsi residual yang harus dipenuhi dalam pemodelan regresi linear yaitu residual bersifat identik, independen, dan berdistribusi normal.

a. Uji Identik

Pengujian identik menggunakan Uji *Glejser* bertujuan untuk melihat apakah ada kesamaan varian dari residual untuk semua pengamatan pada model regresi linear. Berdasarkan hasil Uji *Glejser* dengan taraf signifikansi (α) 0,05 didapatkan nilai *p-value* sebesar 1,000. Hal tersebut menandakan varian dari residual identik karena *p-value* lebih besar dari taraf signifikansi.

b. Uji Independen

Uji independen digunakan untuk mengetahui apakah residual sudah memenuhi asumsi independen atau tidak. Pengujian dapat menggunakan uji *Durbin-Watson*. Dimana, berdasarkan hasil perhitungan didapatkan nilai *Durbin-Watson* adalah sebesar 2,01798. Dimana dengan taraf signifikansi 5%, nilai DU dan 4-DL adalah sebesar 1,33 dan 2,52, maka dapat diputuskan gagal tolak H_0 . Sehingga dapat disimpulkan bahwa residual indenpenden karena nilai *Durbin-Watson* berada pada interval $d_U < d < 4 - d_L$

c. Uji Distribusi Normal

Pengujian distribusi normal salah satunya dapat menggunakan uji *Kolmogorov-Smirnov*. Pengujian distribusi normal digunakan untuk melihat apakah residual bersifat normal atau tidak. Dari hasil perhitungan didapatkan nilai KS sebesar 0,211 atau nilai *p-value* lebih kecil dari 0,01. Hal tersebut menandakan bahwa residual tidak berdistribusi normal pada taraf signifikansi 5%, karena *p-value* yang didapatkan lebih kecil dari taraf signifikansi.

4.3.5 Model Regresi Linear Berganda pada Jam Tidur Tidak Nyenyak Pasien

Pada pemodelan variabel rata-rata Jam tidur tidak nyenyak pasien setiap harinya selama 14 hari setelah melakukan kemoterapi didapatkan tabel ANOVA yang ditunjukkan pada Tabel 4.10.

Tabel 4.10 *Analysis of Variance* (ANOVA) variabel Y_3

<i>Source</i>	<i>DF</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>P</i>
<i>Regression</i>	2	30139	15069	3,18	0,059
<i>Error</i>	25	118510	4740		
<i>Total</i>	27	148649			

Tabel 4.10 merupakan hasil *Analysis of Variance* (ANOVA) yang memberikan hasil bahwa minimal ada satu variabel yang berpengaruh signifikan terhadap jam tidur tidak nyaman pasien setelah melakukan kemoterapi pada taraf signifikansi (α) sebesar 0,1. hal tersebut dapat dilihat dari nilai *p-value* sebesar 0,059 dimana nilai tersebut kurang dari taraf signifikansi (α).

Setelah mengetahui bahwa terdapat minimal satu variabel yang berpengaruh signifikan terhadap jam tidur tidak nyaman pasien setelah melakukan kemoterapi, kemudian dilakukan pengujian parsial untuk mengetahui variabel mana saja yang berpengaruh signifikan terhadap jam tidur tidak nyaman pasien setelah melakukan kemoterapi.

Tabel 4.11 Estimasi dan Pengujian Parameter Model Regresi Linear Y_3

Prediktor	Coef	SE	T	P	VIF
Constant	88,6	81,5	1,09	0,288	
X_2	2,30	1,27	1,82	0,081	1,12
X_{12}	-0,2199	0,098	-2,24	0,034	1,12

Tabel 4.11 menunjukkan bahwa nilai VIF tidak lebih dari 10. Hal ini menunjukkan bahwa pada model regresi menggunakan metode *stepwise* tidak terdapat multikolinearitas antara variabel prediktor yang dimasukkan kedalam model. Kemudian pada pengujian signifikansi variabel dengan taraf signifikansi (α) sebesar 0,05 didapatkan bahwa variabel dosis senyawa P berpengaruh signifikan terhadap jam tidur tidak nyaman pasien setelah melakukan kemoterapi dan variabel usia signifikan terhadap jam tidur pasien pada taraf signifikansi 0,1. Kemudian akan dibentuk model persamaan regresi linear berganda yang terbentuk untuk mengetahui besarnya pengaruh variabel usia dan dosis obat senyawa P terhadap jam tidur tidak nyaman pasien setelah melakukan kemoterapi menggunakan metode *stepwise*. Berikut adalah model persamaan regresi linear yang didapatkan.

$$\hat{Y}_3 = -1489 + 2,30 X_2 - 0,2199 X_{12}$$

Keterangan :

X_2 = Usia pasien

X_{12} = dosis senyawa P

Persamaan tersebut menjelaskan bahwa setiap peningkatan satu satuan pada usia pasien dan variabel lainnya dianggap konstan, maka akan mengakibatkan peningkatan jam tidur tidak nyenyak pasien sebesar 2,3 menit. Sedangkan pada dosis obat senyawa P meningkat satu satuan akan mengakibatkan penurunan pada jam tidur tidak nyenyak sebesar 0,2199 menit.

Model regresi linear diatas menghasilkan nilai koefisien determinasi (R^2) sebesar 20,28% yang artinya bahwa variabel usia pasien dan dosis obat senyawa P dapat menjelaskan jam tidur tidak nyenyak pasien sebesar 20,28%, sedangkan sisanya dijelaskan oleh variabel lain yang tidak dimasukkan kedalam model.

4.3.6 Pengujian Asumsi Model Regresi Linear Jam tidur

Tidak Nyenyak Pasien

Terdapat asumsi residual yang harus dipenuhi dalam pemodelan regresi linear yaitu residual bersifat identik, independen, dan berdistribusi normal.

a. Uji Identik

Berdasarkan hasil perhitungan didapatkan hasil dari Uji *Glejser* nilai *p-value* sebesar 1,00. Dengan menggunakan taraf signifikansi 5% dapat disimpulkan bahwa varian dari residual sudah sama atau sudah memenuhi asumsi identik.

b. Uji Independen

Pada pengujian independen dapat dilakukan menggunakan uji *Durbin-Watson*. Berdasarkan hasil perhitungan, didapatkan nilai *Durbin-Watson* sebesar 1,86913. dimana nilai D_U dan $4-D_L$ adalah sebesar 1,26 dan 2,44 pada taraf signifikansi 5%. Maka dapat diambil keputusan gagal tolak H_0 dan disimpulkan bahwa residual saling indenpenden karena nilai *Durbin-Watson* berada pada interval $d_U < d < 4 - d_L$.

c. Uji Distribusi Normal

Pengujian distribusi normal ingin diketahui apakah residual dari model regresi bersifat normal atau tidak. Pengujian normal dapat dilakukan menggunakan uji *Kolmogorov-Smirnov*. Dimana berdasarkan hasil perhitungan didapatkan nilai KS sebesar 0,143 atau *p-value* sebesar 0,147. sehingga dapat disimpulkan bahwa

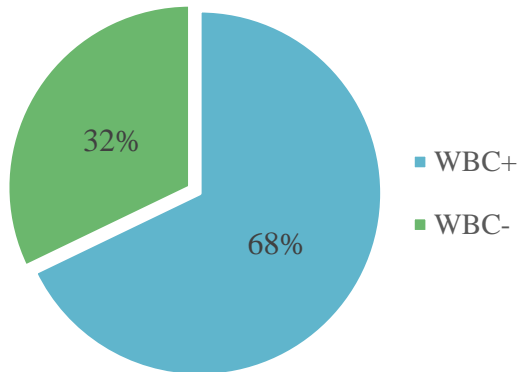
residual dari model sudah berdistribusi normal karena p -value lebih besar dari taraf signifikansi.

4.4 Pemodelan menggunakan Analisis Regresi Logistik

Setelah melakukan pemodelan menggunakan analisis regresi multivariat dan regresi linear berganda, selanjutnya pada data kasus kemoterapi kanker berbasis 5-FU akan dilakukan pemodelan menggunakan analisis regresi logistik biner. Pada analisis regresi logistik biner dilakukan untuk mengetahui variabel prediktor yang berpengaruh terhadap pengklasifikasian keberhasilan kemoterapi kanker.

4.4.1 Karakteristik Variabel yang Digunakan pada Analisis Regresi Logistik Biner

Pada sub bab ini akan dilakukan analisis regresi logistik biner untuk mengetahui faktor-faktor apa saja yang mempengaruhi peningkatan atau penurunan jumlah sel darah putih (WBC). Variabel respon yang digunakan akan diklasifikasikan menjadi 2 kategori yaitu jumlah WBC pasien mengalami peningkatan setelah melakukan kemoterapi dan WBC pasien mengalami penurunan jumlah setelah melakukan kemoterapi. Proporsi masing-masing ditunjukkan pada Gambar 4.10



Gambar 4.10 Proporsi Klasifikasi Perubahan WBC pada Pasien

Proporsi perubahan pada jumlah total *White Blood Cells* yang mengalami peningkatan dan penurunan tidak *balance* seperti yang dapat dilihat pada Gambar 4.10. Dimana perbandingan jumlah total

White Blood Cells pada pasien yang mengalami peningkatan dan penurunan setelah melakukan kemoterapi adalah masing-masingnya adalah 79% dengan 21%.

4.4.2 Estimasi Parameter Menggunakan Metode Wald

Pada pengujian analisis regresi logistik biner untuk mendapatkan model terbaik maka perlu dilakukan seleksi model terbaik agar semua parameter yang dimasukkan kedalam model signifikan. Metode yang digunakan adalah metode *Backward Stepwise (Wald)*, yaitu memasukkan variabel yang signifikan dan mengeluarkan variabel yang tidak signifikan dari model dengan tujuan untuk mendapatkan model yang terbaik.

Hasil pengujian signifikansi menggunakan metode *Backward Stepwise (Wald)* untuk mendapatkan model terbaik ditunjukkan oleh Tabel 4.11

Tabel 4.12 Seleksi Model Terbaik

Variabel	B	p-value	OR	95% C.I for OR	
				Lower	Upper
X ₂	-0,275	0,059	0,759	0,570	1,011
X ₄	-0,001	0,034	0,999	0,997	1,000
X ₅	-0,026	0,073	0,974	0,947	1,002
Konstanta	44,242	0,046	1,637x10 ¹⁶		

Tabel 4.12 menjelaskan bahwa usia pasien, rata-rata jumlah langkah kaki pasien selama 14 hari dan jumlah dosis 5 FU-1 yang diberikan pada saat kemoterapi berpengaruh signifikan pada perubahan WBC setelah melakukan kemoterapi. Berdasarkan nilai *odds ratio* yang didapatkan jika semakin meningkat umur pasien, maka kemungkinan pasien mengalami peningkatan pada WBC setelah melakukan kemoterapi lebih besar 0,759 kali. begitu juga dengan variabel lainnya yang masuk ke dalam model.

Model terbaik yang dapat menjelaskan perubahan WBC setelah melakukan kemoterapi adalah sebagai berikut.

$$\pi_i(\mathbf{x}_i) = \frac{\exp\left(\frac{\mu(\mathbf{x})}{s(\mathbf{x})}\right)}{1 + \exp\left(\frac{\mu(\mathbf{x})}{s(\mathbf{x})}\right)}$$

$$1 - \pi_i(\mathbf{x}_i) = \frac{1}{1 + \exp(\mathcal{G}(x))}$$

Dimana $\mathcal{G}(x) = 44,242 - 0,026x_2 - 0,001x_4 - 0,275x_5$

Maka dapat disimpulkan pada model $\pi_i(\mathbf{x}_i)$ dapat digunakan untuk menghitung peluang pasien mengalami peningkatan pada WBC setelah melakukan kemoterapi jika umur pasien semakin bertambah, rata-rata jumlah langkah kaki dan dosis senyawa 5-*Fluorouracil* 1 (5-FU 1) yaitu sebesar $\pi_i(\mathbf{x}_i)$. Sedangkan peluang seorang pasien mengalami penurunan WBC setelah melakukan kemoterapi jika umur pasien semakin bertambah, rata-rata jumlah langkah kaki dan dosis senyawa 5-FU 1 yaitu sebesar $1 - \pi_i(\mathbf{x}_i)$.

4.4.3 Ketepatan Klasifikasi

Untuk mengukur tingkat akurasi dari hasil klasifikasi dan model yang terbentuk, maka akan dihitung nilai ketepatan klasifikasi dan nilai akurasinya. Tabel ketepatan klasifikasi antara nilai observasi dan hasil prediksi model akan ditampilkan pada Tabel 4.13.

Tabel 4.13 Hasil Prediksi Klasifikasi WBC

Observasi		Prediksi	
		Perubahan WBC	
Perubahan WBC	Menurun	Menurun	Meningkat
	Meningkat	3	16

Tabel 4.13 merupakan hasil tabulasi silang hasil klasifikasi perubahan WBC setelah melakukan kemoterapi. Nilai ketepatan klasifikasi dapat menggunakan nilai AUC yang dihitung menggunakan Persamaan (2.43). Dimana, dari hasil perhitungan didapatkan nilai AUC sebesar 75,44%. sehingga dapat disimpulkan bahwa hasil pengklasifikasian dan model yang terbentuk cukup baik untuk menjelaskan perubahan WBC setelah melakukan kemoterapi berbasis 5-FU pada pasien dengan tingkat ketepatan klasifikasi 75,44%.

4.4.4 Hasil Analisis Regresi Logistik Biner

Kemoterapi yang digunakan pada penelitian ini memberikan *treatment* yang relatif baru pada dunia kemoterapi. Dalam penelitian ini mencoba untuk menganalisis lebih jauh mengenai efek kemoterapi berbasis *5-Fluorouracil* yang dikombinasikan dengan *treatment* lain, yaitu dengan *activity meter* yang dipasangkan pada tubuh pasien setelah melakukan kemoterapi. Selama 14 hari setelah menjalani kemoterapi, *activity meter* digunakan untuk mengukur jumlah langkah kaki dan jam tidur tidak nyenyak pada pasien. Berdasarkan hasil pengujian Regresi logistik biner didapatkan hasil bahwa usia pasien, rata-rata jumlah langkah kaki dan dosis senyawa *5-Fluorouracil* 1 berpengaruh signifikan terhadap efek samping kemoterapi.

Pada hasil pengujian tersebut terdapat 6 jenis senyawa kemoterapi yang tidak berpengaruh signifikan terhadap model yang didapatkan, yaitu senyawa *levamisole*, *oxaliplatin*, *irinotecan*, *Bevacizumab*, *capecitabine* dan *panitumumab*. 6 jenis senyawa tersebut merupakan senyawa anti kanker yang memiliki efek yaitu terjadinya penurunan pada *total white blood cells count*. Namun, pada penelitian ini didapatkan bahwa senyawa tersebut tidak memberikan pengaruh signifikan terhadap penurunan *total white blood cells count*. Pada penelitian sebelumnya didapatkan hasil kesimpulan yang sama bahwa senyawa tersebut tidak memberikan pengaruh terhadap menurunnya *total white blood cell count* seperti pada penelitian yang dilakukan oleh Chai, et al (2011) mendapatkan hasil bahwa tidak terdapat perbedaan yang signifikan dari nilai *total white blood cell count (WBC)*, *absolute neutrophil count (ANC)* dan *absolute lymphocyte count (ALC)* antara pasien yang menggunakan senyawa *levamisole* dan pasien yang tidak menggunakan senyawa *levamisole*. Pada penelitian lainnya oleh Si, W et al (2013) juga mendapatkan hasil bahwa tidak

ada perbedaan yang signifikan pada pasien melakukan kemoterapi yang diberi senyawa *capecitabine* dengan pasien yang tidak diberi senyawa *capecitabine* terhadap *Gastrointestinal*, penurunan jumlah sel darah putih (WBC) dan *palmar-plantar erythrodysesthesia* yang merupakan efek samping utama dari pemberian kemoterapi yang dikombinasikan dengan *capecitabine*.

BAB V

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Kesimpulan yang didapatkan berdasarkan analisis yang telah dilakukan adalah sebagai berikut:

1. Berikut ini adalah ringkasan karakteristik faktor-faktor yang diduga berpengaruh terhadap perubahan *White Blood Cells*, *neutrophil* dan rata-rata jam tidur tidak nyenyak pasien, setelah melakukan kemoterapi
 - a. Pasien berjenis kelamin laki laki adalah sebanyak 79% dari total pasien dan jumlah pasien kemoterapi kanker yang berjenis kelamin perempuan adalah sebanyak 21%.
 - b. Pasien yang melakukan kemoterapi memiliki nilai indeks massa tubuh 18,5 – 24,9 adalah sebanyak 53% dari total pasien Sedangkan sisanya, 11 % dari total pasien memiliki IMT kurang dari 18,5. Dan masing-masing 18 % memiliki IMT sebesar 25,0 - 29,9 dan 30 – 39,9.
 - c. Rata-rata usia pasien kemoterapi kanker memiliki usia 64,96 tahun, dengan standar deviasi 11,06. umur minimum dari pasien kemoterapi adalah 39 tahun dan umur *maximum* 84 tahun.
 - d. Rata-rata jumlah langkah kaki pasien adalah 5333 langkah perhari.
 - e. Rata-rata dosis yang digunakan untuk senyawa kemoterapi kanker 5-FU 1 adalah 641,6. Dimana dosis *maximum* yang didapat pasien adalah masing-masingnya 750 dan 450
 - f. Rata-rata dosis 5-FU 2 adalah 3823,2 dan 327,5. Dimana dosis *maximum* yang diterima pasien adalah 4650 dan dosis minimumnya adalah 2800
 - g. Rata-rata dosis senyawa *Levamisole* (LV) yang diterima pasien adalah 327,5 dengan dosis *maximum* 380 dan dosis minimumnya adalah 250.
 - h. Senyawa lain yang digunakan dalam kemoterapi kanker adalah *Oxaliplatin* (L-OHP), *Irinotecan* (IRI), *Bevacizumab* (B), *Capecetabine* (C), dan *Panitumumab*

(P) dimana pada kasus kemoterapi ini, persentase pasien yang menggunakan senyawa tersebut masing-masingnya adalah 62%, 7%, 24%, 17% dan 17%.

2. Analisis regresi multivariat mendapatkan kesimpulan bahwa tidak terdapat variabel yang mempengaruhi perubahan *White Blood Cells*, perubahan nilai total *neutrophil* dan rata-rata jam tidur tidak nyenyak pasien setelah melakukan kemoterapi. Setelah dilakukan analisis secara univariat menggunakan regresi linear berganda dengan metode *stepwise* didapatkan bahwa terdapat dua variabel respon yang berpengaruh signifikan terhadap perubahan *White Blood Cells* setelah melakukan kemoterapi yaitu dosis senyawa *5-FU* 1 dan Indeks Massa Tubuh pasien pada taraf signifikansi 5% sedangkan variabel dosis senyawa *IRI (Irinotecan)* berpengaruh signifikan pada taraf signifikansi sebesar 20%. Kemudian pada perubahan *Neutrophil* setelah melakukan kemoterapi dipengaruhi oleh dosis senyawa *Irinotecan (IRI)* dengan taraf signifikansi 10%. Sedangkan jam tidur tidak nyenyak pada pasien setelah melakukan kemoterapi, dipengaruhi oleh dosis senyawa *Panitumumab (P)* dengan taraf signifikansi 5% dan usia pasien pada taraf signifikansi 10%.
3. Indikator yang berpengaruh signifikan terhadap klasifikasi perubahan *White Blood Cells (WBC)* adalah usia pasien, rata-rata jumlah langkah kaki dan dosis senyawa *5-FU* 1. Kemudian, ketepatan klasifikasi perubahan *White Blood Cells* setelah melakukan kemoterapi adalah sebesar 75,44%.

5.2 Saran

Saran yang dapat diberikan dan digunakan sebagai pertimbangan pada penelitian selanjutnya adalah sebagai berikut:

1. Jam tidur yang tidak nyenyak pada pasien setelah melakukan kemoterapi sebagai salah satu efek samping dari kemoterapi, untuk menekan efek samping tersebut dapat ditekan menggunakan senyawa *panitumumab (P)* karena pada hasil pengujian, didapatkan bahwa senyawa tersebut mampu

mengurangi jam tidur tidak nyenyak pasien setelah melakukan kemoterapi.

2. Penderita kanker yang ingin melakukan kemoterapi disarankan melakukan kemoterapi menggunakan senyawa 5-*Fluorouracil* (5-FU) dikombinasikan dengan senyawa *irinotecan* (IRI) untuk menurunkan jumlah sel darah putih dan jumlah *neutrophil* pada penderita kanker. Selain itu untuk menurunkan jumlah sel darah putih (WBC) pasien juga disarankan untuk membiasakan melakukan jalan kaki karena pada hasil pengujian, didapatkan bahwa jumlah langkah kaki dapat mempengaruhi penurunan jumlah sel darah putih (WBC).
3. Saran untuk penelitian selanjutnya, penelitian yang menggunakan metode Regresi Multivariat apabila tidak ditemukan variabel yang mempengaruhi model multivariat, penulis menyarankan untuk melakukan analisis menggunakan regresi linear untuk melihat pengaruh terhadap masing-masing respon yang digunakan. Penelitian selanjutnya menggunakan analisis regresi logistik biner, disarankan menggunakan data yang *balance* dan juga sampel yang mencukupi agar mendapatkan hasil analisis yang lebih baik

(Halaman ini sengaja dikosongkan)

DAFTAR PUSTAKA

- Agresti, A. (2002). *Categorical Data Analysis* (3rd ed.). New York: John Wiley & Sons, Inc.
- Agresti, A. (2007). *An Introduction to Categorical Data Analysis Second Edition*. United State of America: A John Wiley & Sons, Inc.
- Bekkar, M., Djemaa, H. K. & Alitouche T. A. (2013). *Evaluation Measures for Models Assessment over Imbalanced Data Sets*. J. Inf. Eng. Appl. 3, 27–38.
- Cameron, D. (2009). Management Of Chemotherapy-Associated Febrile Neutropenia. *British Journal of Cancer*, vol 101, 18-22.
- Chai, P. R., Bastan, W., Machan, J., Hack, J. B. & Babu, K. M. (2011). Levamisole exposure and hematologic indices in cocaine users. *Official Journal of the Society for Academic Emergency*, 1-7.
- Chapuis, P. H., Dent, O. F., Fisher, R. Newland, R. C., Pheils, M. T., Smyth, E. & Colquhoun, K. (1985). A multivariate analysis of clinical and pathological variabels in prognosis after resection of large bowel cancer. *Br. J. Surg*, Vol 72, 698-702 .
- Cancer Information Service. (2017). *Projected Cancer Statistics, 2017*. Dipetik Februari 18, 2018, dari https://ganjoho.jp/en/public/statistics/short_pred.html
- Desen, W. (2008). *Onkologi Klinis (Edisi 2)*. Jakarta: Balai Penerbit Fakultas Kedokteran Universitas Indonesia.
- Drapper, N. R. & Smith, H.. (1991). *Analisis Regresi Terapan*. Jakarta: PT. Gramedia Pustaka Utama.
- Fortheringham, A. S., Brudson, C., & Charlton, M. E. (2002). *Geographilly Weghited Regression: The Analysis of Spatially Varying Relationship*. England: John Wiley and Sons LTd.

- Gale, D. (2000). *Rencana Asuhan Keperawatan Onkologi*. Jakarta: EGC.
- Giacchetti, S., Itzhaki, M., Gruia, G., Adam, R., Zidani, R., Kunstlinger, F., et al. (1999). Long-term survival of patients with unresectable colorectal cancer liver metastases following infusional chemotherapy with 5-Fluorouracil, leucovorin, oxaliplatin and surgery. *Annals of Oncology*, Vol 10, 663–669.
- Gorunescu, F. (2011). *Data Mining: Concepts, Models and Techniques*. Berlin: Springer-Verlag Berlin Heidelberg.
- Gujarti, D. (2003). *Basic Econometrics*. (S. Zain, Trans.) Jakarta: Erlangga.
- Hastie, T., Tibshirani, R., & Friedman, J. (2001). *The Elements of Statistical Learning*. Springer Verlag.
- Heather, L. & Gerald, R. D. (2008) Outpatient management of febrile neutropenia: concern for the future. *J Support Oncol*, Vol 6, 217- 218.
- Hocking, R.R. (2003). *Methods and Applications of Linear Models 2nd Edition*. New Jersey: John Wiley & Sons.
- Hosmer, D.W., & Lemeshow. (2000). *Applied Logistik Regression*. USA : John Wiley and Sons.
- Hosmer, D.W., Lemeshow, S. & Sturdivant, X.R. (2013). *Applied Logistik Regression 3rd Edition*. New Jersey: John Wiley & Sons.
- Japan Cancer Society. (2011). *Cancer in Japan*. Dipetik Februari 18, 2018, dari <http://www.jcancer.jp/en/cancer-in-japan>
- Johnson, R.A. & Wichern, D.W. (2007). *Applied Multivariate Statistical Analysis*, 6th edition. New Jersey: Printice Hall.
- Kartika, U. (2013). *Penderita Kanker di Indonesia Meningkat*. Dipetik Februari 10, 2018, dari <http://www.health.kompas.com>
- Kiswari, R. (2014). *Hematologi & Transfusi*. Jakarta: Erlangga.

- Kohne, C.H., Cunningham, D., Costanzo, F.D., Glimelius, B., Blijham, G., Aranda, E., et al (2002). Clinical determinants of survival in patients with 5-Fluorouracil-based treatment for metastatic colorectal cancer: results of a multivariate analysis of 3825 patients. *Annals of onkology*, Vol 13, 308-317
- Kumar, V., Cotran, R.S. & Robbins, S.L. (2007). *Buku Ajar Patologi: Bab 6 Neoplasma*. Ed 7. Jakarta: EGC, 186- 213.
- Longley, D.B. & Johnston, P.G. (2007). *5-Fluorouracil Molecular Mechanisms of Cell Death in Srivastava R., Apoptosis, Cell Signaling, and Human Diseases*. New York: Humana Press.
- Maalouf, M. & Trafalis, T.B. (2010). Robust weighted kernel logistik regression in imbalanced and rare. *Computational Statistics and Data Analysis 2011*, Vol 55, 168-183.
- Morrison, D.F. (2005). *Multivariat Statistical Methods Fourth Edition*. The Philadelphia: Wharton School University of Pennsylvania.
- Meyerhardt, J.A. & Mayer, R.J., (2005). Systemic Therapy for Colorectal Cancer, *N. Engl. J. Med.*, 476-487.
- Moertel, C.G., Thomas, M.D., Fleming. T.R., Macdonald. J.S., Haller. D. G., Laurie. J.A., Tangen. C.M., et al. (1995). Fluorouracil Plus Levamisole as Effective Adjuvant Therapy After Resection of Stage III Colon Carcinoma. *Annals of Internal Medicine*, Vol 122, 321-326.
- National Cancer Institute (2015). *Chemotherapy*. Dipetik Februari 17, 2018, dari: <https://www.cancer.gov/about-cancer/treatment/types/chemotherapy>
- Otto, S.E. (2003). *Buku Saku Keperawatan Onkologi*. Jakarta: EGC.

- Peng, J., Zhu, J., Bergamaschi, A., Han, W., Noh, D. Y., Pollack, J. R., & Wang, P. (2014). Regularized Multivariate Regression for Identifying Master Predictors with Application to Integrative Genomics Study of Breast Cancer. *The Annals of Applied Statistics*, Vol 4, 53-77.
- Price. (2006). *Patofisiologi Konsep Klinis Proses-proses Penyakit Ed 6*. Jakarta: EGC, 150-158.
- Rencher, A. C. (2002). *Methods of Multivariate Analysis*. Canada: John Wiley & Sons, Inc.
- Sadikin, M. (2002). *Biokimia Darah*. Jakarta. Widiya Medika.
- Setiarni, S. M., Sutomo, A. H. & Hariyono, W. (2011). Hubungan Antara Tingkat Pengetahuan, Status Ekonomi Dan Kebiasaan Merokok Dengan Kejadian Tuberkulosis Paru Pada Orang Dewasa Di Wilayah Kerja Puskesmas Tuan-Tuan Kabupaten Ketapang Kalimantan Barat. *KES MAS Vol 5*, 162-232.
- Sharma, A. & Lokeshwar, N. (2005). Febrile neutropenia in haematological malignancies. *J Postgrad Med* , Vol 51, 42-48.
- Si, W., Zhu, Y. Y., Li, Y., Gao, P., Han, C., You, J. H., Linghu, R. X., Jiao, S. C. & Yang, J. L. (2013). Capecitabine Maintenance Therapy in Patients with Recurrent or Metastatic Breast Cancer. *Brazilian Journal of Medical and Biological Research*, Vol 46, 1074-1081.
- Smeltzer, S.C. (2002). *Buku Ajar Keperawatan Medikal Bedah (Edisi 8)*. Jakarta: EGC.
- Sutedjo, A. (2006). *Buku Saku Mengenal Penyakit Melalui Pemeriksaan Hasil Laboratorium Edisi Revisi*. Yogyakarta : Amara Books.
- Tamoki, M. (2017). *Extraction of side effect characteristics given by 5-FU using machine learning*.

- Thomas, J.B.D., Sharker, A. & Glenne-Jones, R., (2004). Chest Pain Induced by 5-Fluorouracil. *Br. J., Cardiol, Vol 11*, 483-485.
- Vercellis .(2009). *The Business Intelligence*. Italy :Politecnico di milano.
- Wahyuni, T. (2015). Hubungan Antara Frekuensi Kemoterapi Dengan Kualitas Hidup Perempuan Dengan Kanker Payudara Yang Menjalani Kemoterapi Di Ruang Kemoterapi Rsud. A.M Parikesit Tenggarong. *Jurnal Ilmu Kesehatan, Vol 3*, 1-3.
- World Health Organization (2018). *Cancer: Fact Sheets*. Dipetik Februari 10, 2018, dari: <http://www.who.int/media/centre/factsheets/fs297/en/index.html>
- World Health Organization (2007). *BMI Classification*. Dipetik Maret 26, 2018, dari: http://apps.who.int/bmi/index.jsp?introPage=intro_3.html
- Yan, X., & Su, X.G. (2009). *Linear Regression Analysis : Theory and Computing* . Singapore: World Scientific.
- Yayasan Kanker Indonesia. (2018). *Tentang Kanker*. Dipetik Februari 10, 2018, dari: <http://yayasankanker-indonesia.org/tentang-kanker>

(Halaman ini sengaja dikosongkan)

LAMPIRAN

Lampiran 1. Data Pasien Kemoterapi Kanker Berbasis 5-FU

	Y1	Y2	Y3	Y4	X1	X2	X3	X4
1	-300	-376,9	257,86	0	1	63	2	5750,57
2	500	-1167,5	225,29	1	0	46	4	5168,86
3	-400	-373,6	137,00	0	1	48	3	7856,00
4	600	148,2	329,07	1	1	71	2	7586,00
5	800	485,9	288,29	1	1	71	3	1753,29
6	-500	-376,5	346,93	0	1	69	2	5377,29
7	-1100	-896,9	325,15	0	1	64	2	5795,86
8	-900	-511,6	168,73	1	1	59	2	5180,86
9	100	181,5	301,62	0	1	79	2	5538,43
10	-700	-472,2	249,93	0	0	57	1	10136,43
11	-100	-528,8	171,50	1	1	57	2	7059,29
12	100	-321,1	264,29	1	1	75	3	2902,43
13	-1400	-1028	176,00	1	1	74	2	2883,00
14	-1200	-927,7	109,71	0	1	84	1	2089,00
15	1600	873,6	219,14	1	0	64	3	3564,86
16	1100	997	109,62	1	0	72	2	4547,43
17	800	-39,4	184,14	1	0	70	2	3974,14
18	-300	-262,7	340,00	0	1	77	4	5808,00
19	1000	853,2	208,31	1	1	65	2	7016,57
20	-500	-529	280,36	0	1	65	2	10726,00
21	-1600	-1115,2	104,50	1	1	64	2	6883,29
22	2000	1209,5	192,57	1	1	72	3	3996,57
23	-1500	-701,1	244,23	1	0	47	4	5155,71
24	-1300	-1479,3	136,07	1	1	39	1	4112,00
25	800	356,4	204,00	1	1	57	2	6948,14
26	-2200	-1460,8	189,67	1	1	82	2	3045,43
27	1100	1094,3	331,25	1	1	68	4	6345,23
28	800	1103,1	180,92	1	1	60	4	2127,86

Lampiran 1. Data Pasien Kemoterapi Kanker Berbasis *5-FU*
(lanjutan)

X5	X6	X7	X8	X9	X10	X11	X12
730	4400	350	150	0	0	0	0
650	3700	330	0	0	350	0	0
750	4650	350	160	0	0	950	0
500	3600	300	0	0	250	0	0
750	4500	380	160	0	390	0	0
650	3900	330	140	0	0	0	0
650	3800	350	0	250	300	0	0
750	4500	375	150	0	350	0	0
620	3700	310	130	0	0	0	0
560	3350	280	120	0	0	0	0
670	4000	330	140	0	0	800	0
750	4500	370	160	0	0	950	0
660	3900	330	140	0	0	0	0
640	3500	320	0	0	0	0	300
640	3800	320	135	0	0	0	0
560	3350	280	120	0	0	0	0
510	3100	280	110	0	0	0	240
630	3800	320	130	0	0	0	0
450	2800	340	100	0	0	0	0
680	4000	325	140	0	0	0	0
660	4000	330	0	250	290	0	0
700	4300	350	0	0	0	0	430
650	3400	330	0	0	350	0	0
525	3100	250	0	0	0	660	0
650	3500	320	0	0	0	800	0
600	3600	320	0	0	0	0	360
630	3800	320	130	0	0	0	0
750	4500	380	160	0	0	0	450

Keterangan :

- Y1 : Perubahan *White Blood Cells* (WBC) pasien setelah melakukan kemoterapi
- Y2 : Perubahan persentase kandungan *Neutrophil* pada WBC pasien setelah melakukan kemoterapi
- Y3 : Rata-rata jam tidur tidak nyenyak pasien setelah melakukan kemoterapi
- Y4 : Perubahan (naik/turun) *White Blood Cells* (WBC) setelah melakukan kemoterapi
- X1 : Jenis Kelamin Pasien
- X2 : Umur Pasien
- X3 : Indeks Massa Tubuh Pasien
- X4 : Rata-rata jumlah langkah kaki pasien dalam satu hari selama 14 hari setelah melakukan kemoterapi
- X5 : Dosis Senyawa kemoterapi *5-Fluorouracil* (5-FU) 1
- X6 : Dosis Senyawa kemoterapi *5-Fluorouracil* (5-FU) 2
- X7 : Dosis Senyawa kemoterapi *levamisole* (LV)
- X8 : Dosis Senyawa kemoterapi *oxaliplatin* (L-OHP)
- X9 : Dosis Senyawa kemoterapi *Irinotecan* (IRI)
- X10 : Dosis Senyawa kemoterapi *Bevacizumab* (B).
- X11 : Dosis Senyawa kemoterapi *Capecitabine* (C)
- X12 : Dosis Senyawa kemoterapi *Panitumumab* (P)

Lampiran 2. *Output* Minitab Korelasi Antar Variabel

Correlation: Y1; Y2; Y3; X2; X4; X5; X6; X7; X8; X9; X10; X11; X12
Correlations

	Y1	Y2	Y3	X2	X4	X5	X6	X7	X8	X9	X10
Y2	0,884 0,000										
Y3	0,154 0,433	0,163 0,408									
X2	0,070 0,722	0,205 0,296	0,206 0,294								
X4	-0,060 0,760	-0,066 0,737	0,232 0,236	-0,327 0,090							
X5	-0,061 0,759	-0,012 0,951	0,001 0,998	-0,046 0,817	-0,224 0,253						

Lampiran 2. Output Minitab Korelasi Antar Variabel (lanjutan)

	Y1	Y2	Y3	X2	X4	X5	X6	X7	X8	X9	X10
X6	0,045	0,088	0,077	0,041	-0,153	0,922					
	0,821	0,655	0,696	0,838	0,436	0,000					
X7	0,088	0,181	0,134	0,099	-0,254	0,776	0,773				
	0,655	0,357	0,496	0,616	0,192	0,000	0,000				
X8	0,245	0,351	0,184	0,132	0,027	0,352	0,444	0,310			
	0,209	0,067	0,349	0,504	0,893	0,066	0,018	0,109			
X9	-0,330	-0,286	-0,035	-0,025	0,127	0,047	0,045	0,114	-0,364		
	0,086	0,140	0,858	0,901	0,520	0,813	0,820	0,564	0,057		
X10	-0,185	-0,243	0,120	-0,270	-0,047	0,184	0,158	0,324	-0,322	0,415	
	0,347	0,212	0,544	0,165	0,811	0,348	0,422	0,093	0,095	0,028	
X11	-0,019	-0,139	-0,245	-0,372	0,098	0,221	0,196	0,024	0,080	-0,128	-0,263
	0,923	0,481	0,209	0,051	0,618	0,258	0,318	0,903	0,687	0,516	0,176

Lampiran 2. Output Minitab Korelasi Antar Variabel (lanjutan)

	Y1	Y2	Y3	X2	X4	X5	X6	X7	X8	X9	X10
X12	0,104	0,173	-0,312	0,324	-0,478	0,086	0,086	0,142	-0,223	-0,126	-0,258
	0,597	0,380	0,106	0,092	0,010	0,664	0,665	0,472	0,254	0,524	0,184

Lampiran 2. Output Minitab Korelasi Antar Variabel (lanjutan)

	X11
X12	-0,209
	0,286

Cell Contents

Pearson correlation

P-value

Lampiran 3. Macro untuk Pengujian Distribusi Normal Multivariat

```

Macro
NormalMultivariate X.1-X.p qc dj22

MConstant i j n p Prop Tengah
MColumn x.1-x.p xj Kali d dj2 qc Prob dj22
MMatrix MCova MCovaI xjxbar

#-- 1.1. Dapatkan Nilai dj2 --#
let n=count(x.1)
Covariance X.1-X.p MCova
print MCova
invers MCova MCovaI
do i=1:n
  do j=1:p
    let xj(j)=x.j(i)-mean(x.j)
  enddo

  copy xj xjxbar
  mult MCovaI xjxbar Kali
  let d=Kali*xj
  let dj2(i)=sum(d)
enddo
print dj2

#-- 1.2. Dapatkan Nilai qc --#
do i=1:n
  let Prob(i)=1-(n-i+0.5)/n
enddo
INVCDF Prob qc;
Chisquare p.

#-- 1.3 Buat Plot dj2 dengan qc --#
sort dj2 dj22
plot dj22*qc;
symbol.

```

```

#-- 2. Mencari Proporsi --#
INVCDF 0.5 Tengah;
Chisquare p.
let Prop=0
do i=1:n
  if dj2(i)<=Tengah
    let Prop=Prop+1
  endif
enddo
let Prop=Prop/n
print Prop

name qc 'qc'
name dj22 'dj2'
endmacro

```

Lampiran 4. *Output SPSS Uji Independensi Variabel Respon*

KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.516
Bartlett's Test of Sphericity	Approx. Chi-Square	38.892
	df	3
	Sig.	.000

Lampiran 5. Output SPSS Model Regresi Multivariat

Parameter Estimates								
Dependent			Std.			95% Confidence Interval		Partial Eta
Variabel	Parameter	B	Error	t	Sig.	Lower Bound	Upper Bound	Squared
y1	Intercept	2026.273	6026.243	.336	.742	-10992.633	15045.180	.009
	x2	-10.265	29.167	-.352	.731	-73.276	52.746	.009
	x4	.025	.133	.190	.853	-.261	.312	.003
	x5	-8.609	9.696	-.888	.391	-29.556	12.337	.057
	x6	.531	1.865	.285	.780	-3.498	4.561	.006
	x7	6.779	17.527	.387	.705	-31.087	44.644	.011
	x8	.323	6.457	.050	.961	-13.627	14.273	.000
	x9	-2.808	4.021	-.698	.497	-11.494	5.879	.036
	x10	-1.716	3.153	-.544	.595	-8.527	5.095	.022
	x11	-.417	1.034	-.404	.693	-2.650	1.816	.012
	x12	.416	2.755	.151	.882	-5.536	6.369	.002
	X1_0	345.727	929.754	.372	.716	-1662.885	2354.339	.011

y2	X3_1	-1424.948	1014.598	-1.404	.184	-3616.854	766.957	.132
	X3_2	-408.947	670.373	-.610	.552	-1857.201	1039.306	.028
	X3_3	987.962	969.755	1.019	.327	-1107.066	3082.991	.074
	Intercept	1858.728	4508.863	.412	.687	-7882.079	11599.535	.013
	x2	-4.788	21.823	-.219	.830	-51.933	42.357	.004
	x4	.049	.099	.498	.627	-.165	.264	.019
	x5	-3.203	7.254	-.442	.666	-18.876	12.469	.015
	x6	-.359	1.395	-.257	.801	-3.374	2.656	.005
	x7	4.475	13.114	.341	.738	-23.856	32.807	.009
	x8	3.470	4.831	.718	.485	-6.968	13.907	.038
	x9	-.845	3.008	-.281	.783	-7.344	5.655	.006
	x10	-1,000	2.359	-.424	.678	-6.096	4.096	.014
	x11	-.598	.773	-.774	.453	-2.269	1.072	.044
	x12	1.142	2.062	.554	.589	-3.311	5.596	.023
	X1_0	-157.064	695.647	-.226	.825	-1659.917	1345.789	.004
	X3_1	-1105.420	759.127	-1.456	.169	-2745.414	534.574	.140
	X3_2	-389.422	501.576	-.776	.451	-1473.012	694.168	.044

y3	X3_3	722.104	725.575	.995	.338	-845.406	2289.614	.071
	Intercept	164.796	420.000	.392	.701	-742.560	1072.152	.012
	x2	2.298	2.033	1.130	.279	-2.094	6.690	.090
	x4	.009	.009	1.016	.328	-.011	.029	.074
	x5	-.136	.676	-.201	.844	-1.596	1.324	.003
	x6	.009	.130	.067	.948	-.272	.290	.000
	x7	.039	1.222	.032	.975	-2.600	2.678	.000
	x8	-.041	.450	-.090	.930	-1.013	.932	.001
	x9	-.155	.280	-.553	.590	-.760	.451	.023
	x10	.023	.220	.103	.920	-.452	.497	.001
	x11	-.062	.072	-.864	.403	-.218	.093	.054
	x12	-.223	.192	-1.164	.265	-.638	.191	.094
	X1_0	-49.091	64.799	-.758	.462	-189.081	90.900	.042
	X3_1	-96.158	70.713	-1.360	.197	-248.923	56.607	.125
	X3_2	-82.509	46.722	-1.766	.101	-183.445	18.428	.193
	X3_3	-31.862	67.587	-.471	.645	-177.876	114.151	.017

Lampiran 6. Output SPSS Pengujian Signifikansi Mutivariat Model

Multivariate Tests^a

Effect		Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
Intercept	Pillai's Trace	.026	.097 ^b	3.000	11,000	.960	.026
	Wilks' Lambda	.974	.097 ^b	3.000	11,000	.960	.026
	Hotelling's Trace	.027	.097 ^b	3.000	11,000	.960	.026
	Roy's Largest Root	.027	.097 ^b	3.000	11,000	.960	.026
x2	Pillai's Trace	.108	.446 ^b	3.000	11,000	.725	.108
	Wilks' Lambda	.892	.446 ^b	3.000	11,000	.725	.108
	Hotelling's Trace	.122	.446 ^b	3.000	11,000	.725	.108
	Roy's Largest Root	.122	.446 ^b	3.000	11,000	.725	.108
x4	Pillai's Trace	.116	.482 ^b	3.000	11,000	.702	.116
	Wilks' Lambda	.884	.482 ^b	3.000	11,000	.702	.116
	Hotelling's Trace	.131	.482 ^b	3.000	11,000	.702	.116
	Roy's Largest Root	.131	.482 ^b	3.000	11,000	.702	.116
x5	Pillai's Trace	.079	.314 ^b	3.000	11,000	.815	.079
	Wilks' Lambda	.921	.314 ^b	3.000	11,000	.815	.079

x6	Hotelling's Trace	.086	.314 ^b	3.000	11,000	.815	.079
	Roy's Largest Root	.086	.314 ^b	3.000	11,000	.815	.079
	Pillai's Trace	.068	.266 ^b	3.000	11,000	.849	.068
	Wilks' Lambda	.932	.266 ^b	3.000	11,000	.849	.068
	Hotelling's Trace	.073	.266 ^b	3.000	11,000	.849	.068
	Roy's Largest Root	.073	.266 ^b	3.000	11,000	.849	.068
x7	Pillai's Trace	.011	.042 ^b	3.000	11,000	.988	.011
	Wilks' Lambda	.989	.042 ^b	3.000	11,000	.988	.011
	Hotelling's Trace	.012	.042 ^b	3.000	11,000	.988	.011
	Roy's Largest Root	.012	.042 ^b	3.000	11,000	.988	.011
x8	Pillai's Trace	.109	.448 ^b	3.000	11,000	.724	.109
	Wilks' Lambda	.891	.448 ^b	3.000	11,000	.724	.109
	Hotelling's Trace	.122	.448 ^b	3.000	11,000	.724	.109
	Roy's Largest Root	.122	.448 ^b	3.000	11,000	.724	.109
x9	Pillai's Trace	.070	.278 ^b	3.000	11,000	.840	.070
	Wilks' Lambda	.930	.278 ^b	3.000	11,000	.840	.070
	Hotelling's Trace	.076	.278 ^b	3.000	11,000	.840	.070

x10	Roy's Largest Root	.076	.278 ^b	3.000	11,000	.840	.070
	Pillai's Trace	.024	.091 ^b	3.000	11,000	.963	.024
	Wilks' Lambda	.976	.091 ^b	3.000	11,000	.963	.024
	Hotelling's Trace	.025	.091 ^b	3.000	11,000	.963	.024
	Roy's Largest Root	.025	.091 ^b	3.000	11,000	.963	.024
x11	Pillai's Trace	.121	.506 ^b	3.000	11,000	.686	.121
	Wilks' Lambda	.879	.506 ^b	3.000	11,000	.686	.121
	Hotelling's Trace	.138	.506 ^b	3.000	11,000	.686	.121
	Roy's Largest Root	.138	.506 ^b	3.000	11,000	.686	.121
x12	Pillai's Trace	.119	.498 ^b	3.000	11,000	.691	.119
	Wilks' Lambda	.881	.498 ^b	3.000	11,000	.691	.119
	Hotelling's Trace	.136	.498 ^b	3.000	11,000	.691	.119
	Roy's Largest Root	.136	.498 ^b	3.000	11,000	.691	.119
X1_0	Pillai's Trace	.138	.586 ^b	3.000	11,000	.637	.138
	Wilks' Lambda	.862	.586 ^b	3.000	11,000	.637	.138
	Hotelling's Trace	.160	.586 ^b	3.000	11,000	.637	.138
	Roy's Largest Root	.160	.586 ^b	3.000	11,000	.637	.138

X3_1	Pillai's Trace	.241	1.164 ^b	3.000	11,000	.367	.241
	Wilks' Lambda	.759	1.164 ^b	3.000	11,000	.367	.241
	Hotelling's Trace	.317	1.164 ^b	3.000	11,000	.367	.241
	Roy's Largest Root	.317	1.164 ^b	3.000	11,000	.367	.241
X3_2	Pillai's Trace	.235	1.127 ^b	3.000	11,000	.380	.235
	Wilks' Lambda	.765	1.127 ^b	3.000	11,000	.380	.235
	Hotelling's Trace	.307	1.127 ^b	3.000	11,000	.380	.235
	Roy's Largest Root	.307	1.127 ^b	3.000	11,000	.380	.235
X3_3	Pillai's Trace	.094	.383 ^b	3.000	11,000	.768	.094
	Wilks' Lambda	.906	.383 ^b	3.000	11,000	.768	.094
	Hotelling's Trace	.104	.383 ^b	3.000	11,000	.768	.094
	Roy's Largest Root	.104	.383 ^b	3.000	11,000	.768	.094

a. Design: Intercept + x2 + x4 + x5 + x6 + x7 + x8 + x9 + x10 + x11 + x12 + X1_0 + X3_1 + X3_2 + X3_3

b. Exact statistic

Lampiran 7. *Output Minitab Pengujian Signifikasi Parameter Secara Serentak pada Pemodelan Y_1*

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-value
Regression	5	14127419	2825484	3.68	0.014
X5	1	3722365	3722365	4.85	0.038
X9	1	1805760	1805760	2.35	0.139
X3	3	10679067	3559689	4.64	0.012
Error	22	16882224	767374		
Total	27	31009643			

Lampiran 8. *Output Minitab Pengujian Signifikasi Parameter Secara Parsial pada Pemodelan Y_1*

Coefficients

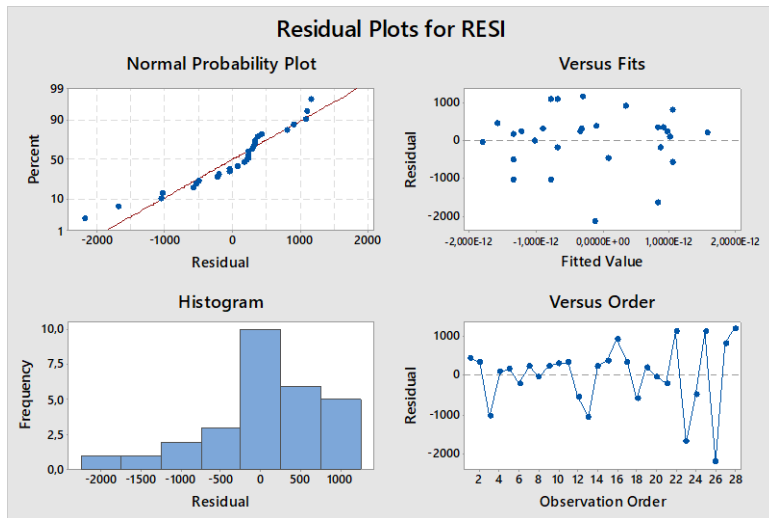
Term	Coef	SE Coef	T-Value	P-value	VIF
Constant	2079	1515	1.37	0.184	
X5	-5.47	2.48	-2.20	0.038	1.42
X9	-4.12	2.69	-1.53	0.139	1.09
X3					
2	1185	571	2.07	0,050	2.96
3	2669	732	3.65	0.001	2.87
4	1663	675	2.46	0.022	2.44

Lampiran 9. koefisien determinasi (R^2) Model regresi linear Y_1

Model Summary

S	R-sq	R-sq(adj)	R-sq(pred)
875,999	45,56%	33,18%	21,13%

Lampiran 10. Output Minitab Asumsi Residual IIDN Model Pada Variabel Y_1

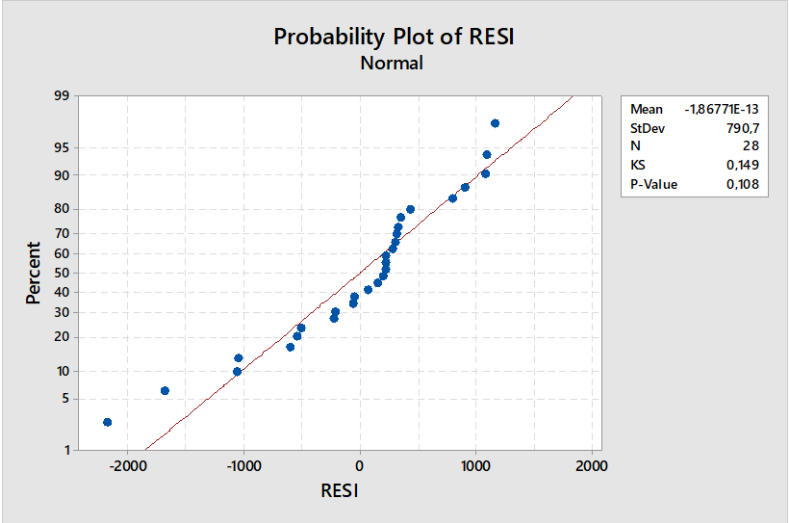


Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	<i>P-value</i>
Regression	5	0	0	0.00	1,000
X5	1	0	0	0.00	1,000
X9	1	0	0	0.00	1,000
X3	3	0	0	0.00	1,000
Error	22	16882224	767374		
Lack-of-Fit	17	12330558	725327	0.80	0.672
Pure Error	5	4551667	910333		
Total	27	16882224			

Durbin-Watson Statistic

Durbin-Watson Statistic = 2.45697



Lampiran 11. *Output Minitab Pengujian Signifikasi Parameter Secara Serentak pada Pemodelan Y_2*

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	<i>P-value</i>
Regression	1	2165629	2165629	3.65	0.067
X8	1	2165629	2165629	3.65	0.067
Error	26	15408115	592620		
Total	27	17573744			

Lampiran 12. *Output Minitab Pengujian Signifikasi Parameter Secara Parsial pada Pemodelan Y_2*

Coefficients

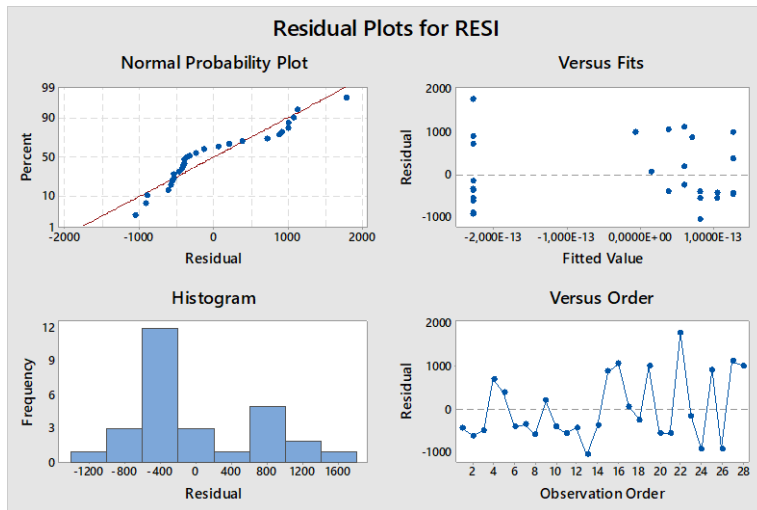
Term	Coef	SE Coef	T-Value	<i>P-value</i>	VIF
Constant	-553	240	-2.30	0.029	
X8	4.13	2.16	1.91	0.067	1.00

Lampiran 13. koefisien determinasi (R^2) Model regresi linear Y_2

Model Summary

S	R-sq	R-sq(adj)	R-sq(pred)
68,8505	20,28%	13,90%	1,53%

Lampiran 14. *Output Minitab Asumsi Residual IIDN Model pada Variabel Y_2*

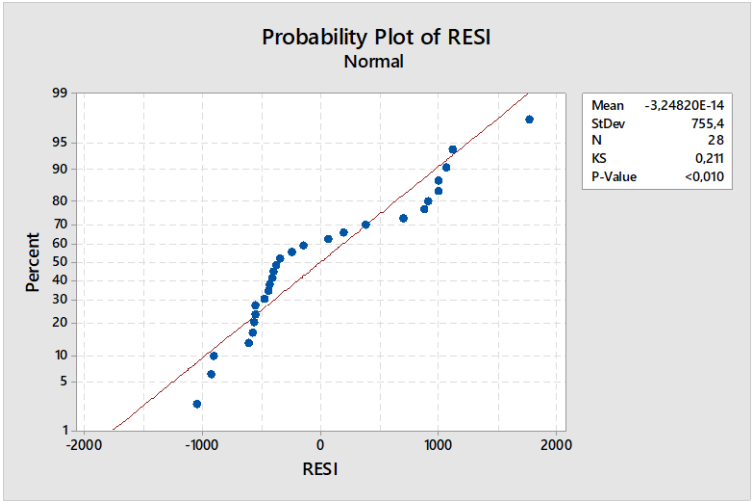


Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	<i>P-value</i>
Regression	1	0	0	0.00	1,000
X8	1	0	0	0.00	1,000
Error	26	15408115	592620		
Lack-of-Fit	7	4568436	652634	1.14	0.378
Pure Error	19	10839679	570509		
Total	27	15408115			

Durbin-Watson Statistic

Durbin-Watson Statistic = 2.01798



Lampiran 15. *Output* Minitab Pengujian Signifikasi Parameter Secara Serentak pada Pemodelan Y_3

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-value
Regression	2	30139	15069	3.18	0,059
X2	1	15654	15654	3.30	0.081
X12	1	23851	23851	5.03	0.034
Error	25	118510	4740		
Total	27	148649			

Lampiran 16. *Output* Minitab Pengujian Signifikasi Parameter Secara Parsial pada Pemodelan Y_3

Coefficients

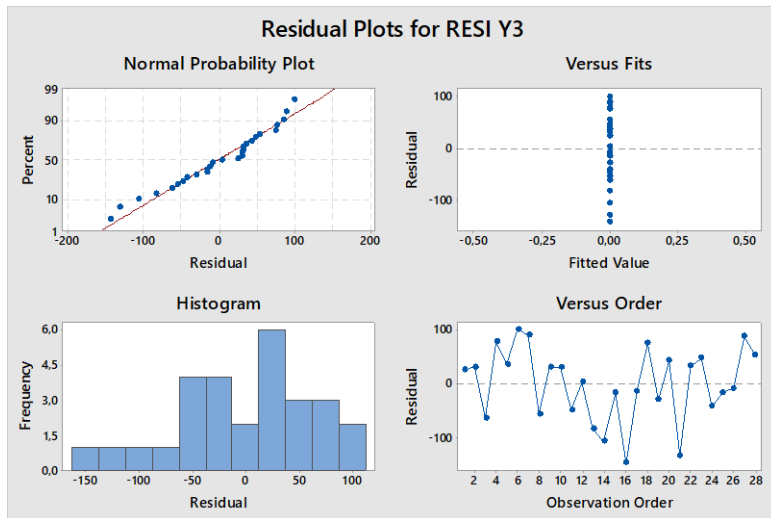
Term	Coef	SE Coef	T-Value	P-value	VIF
Constant	88.6	81.5	1.09	0.288	
X2	2.30	1.27	1.82	0.081	1.12
X12	-0.2199	0.0980	-2.24	0.034	1.12

Lampiran 17. koefisien determinasi (R^2) Model regresi linear Y_3

Model Summary

S	R-sq	R-sq(adj)	R-sq(pred)
769,818	12,32%	8,95%	0,00%

Lampiran 18. Output Minitab Asumsi Residual IIDN Model pada Variabel Y_3

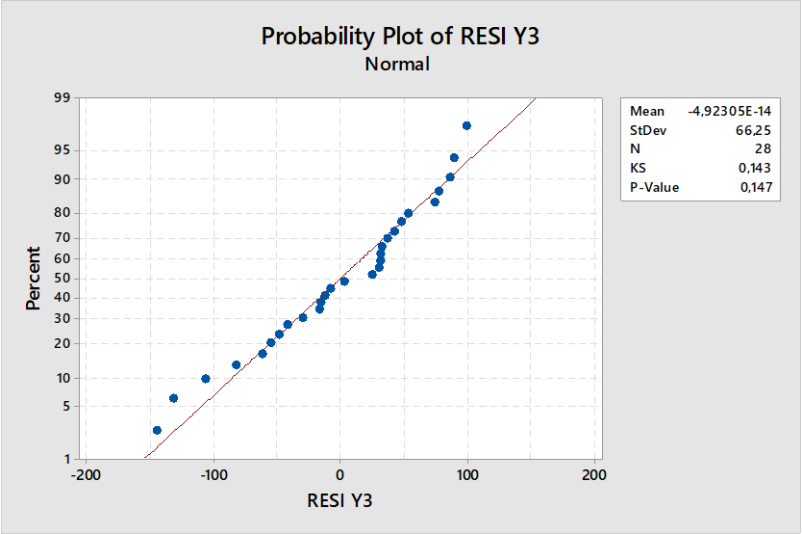


Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-value
Regression	2	0	0.00	0.00	1,000
X2	1	0	0.00	0.00	1,000
X12	1	0	0.00	0.00	1,000
Error	25	118510	4740.40		
Lack-of-Fit	19	87621	4611.61	0.90	0.610
Pure Error	6	30889	5148.23		
Total	27	118510			

Durbin-Watson Statistic

Durbin-Watson Statistic = 1.86130



Lampiran 19. *Output SPSS Model Regresi Logistik Biner Menggunakan Metode Stepwise (Wald)*

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1 ^a	x1(1)	-4.846	64044.772	.000	1	1,000	.008
	x2	-7.990	777.740	.000	1	.992	.000
	x3			.000	3	1,000	
	x3(1)	-61.360	67811.757	.000	1	.999	.000
	x3(2)	-57.590	12100.774	.000	1	.996	.000
	x3(3)	-182.682	54592.069	.000	1	.997	.000
	x4	-.075	6.777	.000	1	.991	.928
	x5	-4.118	623.979	.000	1	.995	.016
	x6	.688	125.593	.000	1	.996	1.989
	x7	-.965	951.082	.000	1	.999	.381
	x8	-1.754	449.915	.000	1	.997	.173
	x9	-.668	142.792	.000	1	.996	.513
	x10	.006	152.980	.000	1	1,000	1.006
	x11	.100	16.696	.000	1	.995	1.105

Step 2 ^a	x12	-.496	145.107	.000	1	.997	.609
	Constant	1553.859	394362.166	.000	1	.997	.
	x1(1)	-4635.694	58476.669	.006	1	.937	.000
	x2	-223.751	2704.775	.007	1	.934	.000
	x4	-1.716	20.744	.007	1	.934	.180
	x5	-96.949	1168.875	.007	1	.934	.000
	x6	10.287	124.124	.007	1	.934	29337.066
	x7	-44.507	540.707	.007	1	.934	.000
	x8	-10.601	129.106	.007	1	.935	.000
	x9	-10.438	125.730	.007	1	.934	.000
	x10	8.813	106.366	.007	1	.934	6720.535
	x11	2.775	33.479	.007	1	.934	16.038
	x12	-3.564	44.261	.006	1	.936	.028
	Constant	63722.946	769654.274	.007	1	.934	.
Step 3 ^a	x2	-225.340	2276.764	.010	1	.921	.000
	x4	-1.727	17.458	.010	1	.921	.178
	x5	-97.856	979.865	.010	1	.920	.000

	x6	10.413	104.222	.010	1	.920	33276.542
	x7	-44.301	459.272	.009	1	.923	.000
	x8	-10.994	109.225	.010	1	.920	.000
	x9	-10.608	105.711	.010	1	.920	.000
	x10	8.786	88.768	.010	1	.921	6540.710
	x11	2.769	27.883	.010	1	.921	15.945
	x12	-3.655	37.085	.010	1	.921	.026
	Constant	63981.719	646716.844	.010	1	.921	.
Step 4 ^a	x2	-.843	.841	1.006	1	.316	.430
	x4	-.006	.006	.998	1	.318	.994
	x5	-.498	.558	.797	1	.372	.608
	x6	.054	.062	.747	1	.388	1.055
	x8	-.096	.121	.627	1	.428	.909
	x9	-.079	.084	.878	1	.349	.924
	x10	.038	.040	.886	1	.347	1.038
	x11	.011	.012	.781	1	.377	1.011
	x12	-.016	.022	.539	1	.463	.984

Step 5 ^a	Constant	219.427	228.775	.920	1	.337	1.977E+95
	x2	-.487	.368	1.753	1	.186	.614
	x4	-.004	.003	1.781	1	.182	.996
	x5	-.324	.289	1.261	1	.261	.723
	x6	.032	.031	1.107	1	.293	1.033
	x8	-.041	.051	.658	1	.417	.960
	x9	-.051	.048	1.142	1	.285	.950
	x10	.038	.032	1.423	1	.233	1.038
	x11	.011	.009	1.549	1	.213	1.011
	Constant	144.247	117.053	1.519	1	.218	4.422E+62
Step 6 ^a	x2	-.266	.188	2.001	1	.157	.766
	x4	-.002	.001	2.999	1	.083	.998
	x5	-.116	.089	1.710	1	.191	.891
	x6	.010	.008	1.513	1	.219	1.010
	x9	-.017	.026	.452	1	.501	.983
	x10	.019	.023	.742	1	.389	1.020
	x11	.005	.004	1.449	1	.229	1.005

Step 7 ^a	Constant	66.816	45.303	2.175	1	.140	10415167767 42332000000 00000000.00 0
	x2	-.271	.185	2.146	1	.143	.762
	x4	-.002	.001	3.850	1	.050	.998
	x5	-.086	.058	2.180	1	.140	.917
	x6	.007	.006	1.642	1	.200	1.007
	x10	.006	.006	1.011	1	.315	1.006
	x11	.003	.003	1.394	1	.238	1.003
	Constant	56.442	33.873	2.776	1	.096	32540761939 33982000000 000.000
Step 8 ^a	x2	-.325	.183	3.161	1	.075	.723
	x4	-.002	.001	4.279	1	.039	.998
	x5	-.073	.046	2.482	1	.115	.929
	x6	.006	.005	1.667	1	.197	1.006

	x11	.002	.002	.806	1	.369	1.002
	Constant	56.711	31.052	3.336	1	.068	42578096000 96133500000 000.000
Step 9 ^a	x2	-.340	.180	3.584	1	.058	.712
	x4	-.002	.001	4.303	1	.038	.998
	x5	-.061	.036	2.773	1	.096	.941
	x6	.005	.004	1.607	1	.205	1.005
	Constant	54.184	28.558	3.600	1	.058	34012661876 43326000000 00.000
Step 10 ^a	x2	-.275	.146	3.554	1	.059	.759
	x4	-.001	.001	4.504	1	.034	.999
	x5	-.026	.014	3.207	1	.073	.974
	Constant	44.242	22.159	3.986	1	.046	16371236342 748072000.0 00

Lampiran 20. *Output* SPSS Ketepatan Klasifikasi

Classification Table^a

		Observed		Predicted		Percentage Correct
			y4			
Step 1	y4	.00	6	3		66.7
		1.00	3	16		84.2
	Overall Percentage					78.6

a. The cut value is .500

Lampiran 21. Surat Pernyataan**SURAT PERNYATAAN**

Saya yang bertanda tangan di bawah ini, mahasiswa Departemen Statistika FMKSD ITS,

Nama : Taufik Afiif Maldini

NRP : 062114 4000 0016

menyatakan bahwa data yang digunakan dalam Tugas Akhir ini merupakan data sekunder yang diambil dari penelitian/ buku/Tugas Akhir/Thesis/publikasi lainnya yaitu :

Sumber : Data Penelitian Tamoki (2017) dengan Judul “*Extraction of side effect characteristics given by 5-FU using machine learning*”

Keterangan : Data merupakan data pasien yang menjalani kemoterapi berbasis 5-Fluorouracil dari *specialist of gastrointestinal surgery* di Nippon Medical School dengan unit observasi sebanyak 28 observasi (pasien), serta jumlah variabel sebanyak 16 variabel.

Surat pernyataan ini dibuat dengan sebenarnya. Apabila terdapat pemalsuan data maka saya siap menerima sanksi sesuai aturan yang berlaku.

Mengetahui,
Pembimbing Tugas Akhir

Surabaya, Juli 2018
Mahasiswa

Dr. rer. pol. Heri Kuswanto, M.Si.
NIP. 19820326 200312 1 004

Taufik Afiif Maldini
NRP.062114 4000 0016

BIODATA PENULIS



Penulis dengan nama lengkap Taufik Afiif Maldini dilahirkan di Kota Bukittinggi, Sumatera Barat pada tanggal 21 Maret 1997. Penulis menempuh pendidikan formal di SDN 02 Percontohan Bukittinggi (2003-2009), SMPN 4 Bukittinggi (2009-2012), dan SMAN 1 Bukittinggi (2012-2014). Kemudian penulis diterima sebagai Mahasiswa Departemen Statistika ITS melalui jalur SNMPTN pada tahun 2014. Selama masa perkuliahan, penulis aktif di organisasi kemahasiswaan ITS tingkat Insitut yakni Badan Eksekutif Mahasiswa ITS sebagai *staff* Badan Eksekutif Mahasiswa ITS pada kementrian BSO ITS Education Care Center (IECC) Pada periode 2015-2016 dan *staff* divisi PERS Himpunan Mahasiswa Statistika ITS periode 2015-2016 kemudian terpilih menjadi wakil ketua Divisi PERS Himpunan Mahasiswa Statistika ITS periode 2016-2017. Selain itu penulis juga aktif dalam berbagai kepanitiaan diantaranya adalah panitia Gerigi ITS 2015, PAMMITS 2015, Pekan Raya Statistika 2016, ITS Mengajar for Indonesia 2016, ITS Expo 2016 dll. Dibidang akademik, penulis pernah mendapatkan juara II pada lomba Infografis NSC Universitas Brawijaya. Apabila pembaca ingin memberi kritik dan saran serta diskusi lebih lanjut mengenai Tugas Akhir ini, dapat menghubungi penulis melalui email taufik.afiif.maldini@gmail.com